Carbohydrate Directed Photoaffinity Labelling

Fowle, Chris

Award date:
2018

Awarding institution:
University of Bath

Carbohydrate Directed Photoaffinity Labelling

Christopher George Fowle

A thesis submitted for the degree of Doctor of Philosophy

University of Bath
Department of Chemistry

September 2017

[*C. G. Fowle*]

Abstract

Glycoproteins have diverse and essential roles within biological systems. They are formed by enzymatic addition of saccharides to proteins during, or shortly after, translation. However, saccharides can also react with proteins non-enzymatically, a process termed glycation, which can cause impaired function and improper folding. Glycated proteins further react to form advanced glycation end-products, which have been implicated in the pathogenesis and progress of many diseases.

Due to this pathological effect, glycation has been studied as a potential biomarker of these diseases. Photoaffinity labelling is a technique that is used to investigate the structure, and presence, of biological molecules; a precedent exists for its use in the study of carbohydrates in biological systems.

Chapter 1 outlines the background of this thesis exploring previous studies of glycation, its effects, and methods used in recognition and photoaffinity labelling.

Chapter 2 details the design and synthesis of a novel photoaffinity probe, and the optimisation of this synthesis. The target molecule was successfully produced and simpler alternatives to the initial synthetic route with similar yields are discussed.

In Chapter 3 the use of the photoaffinity probe is studied. Labelling trials were performed on three proteins: human serum albumin (HSA), macrophage migration inhibitory factor (MIF), and casein. Mass spectrometry showed that the experiments with both HSA and MIF were successful, while the procedure appeared to lead to degradation of casein. Additionally, our work into developing techniques for identifying labelled samples is detailed. A diol-doped electrophoresis gel was not successful created, however, staining protein samples in polyacrylamide gel electrophoresis with curcumin showed promise.

Chapter 4 explores the electrochemistry of the photoaffinity probe and details the use of the probe in functionalising a fluorine doped tin oxide (FTO) glass electrode. Cyclic voltammograms of Alizarin Red S (ARS), obtained using a treated electrode, suggest that surface functionalisation was successful.

Acknowledgements

Firstly, I would like to thank my supervisors, Professor Tony James and Professor Steven Bull for their invaluable advice and guidance. I would also like to thank Tony for his inexhaustible patience, understanding, and support throughout my PhD.

Furthermore, I would like to extend my thanks to those who have helped me with experimental work at various stages. Professor Jean Van Den Elsen and Dr Omar Kassaar Munir for their help with the biochemical aspects of my project. Professor Frank Marken for his assistance with the electrochemical experiments. Also, to all those with whom I shared my time in the James group, especially to Dr Stephen Flower for answering my endless questions during my first year.

Beyond my work, I would like to express my thanks to all the wonderful friends I have been fortunate enough to meet in Bath, for helping to make my time at university as memorable and enjoyable as it has been. While there are many to whom I owe this acknowledgement, I would like to particularly thank Alex, Iain, James, and Tom for their kindness and company, both at the university and in our spare time, during final stages of my PhD, which was greatly appreciated whenever I needed a break from work.

Finally, I offer my thanks to my family for their unconditional love and support, which has helped me to accomplish all that I have.

Abbreviations

3-DG	3-deoxyglucosone
AD	Alzheimer's disease
AGE	advanced glycation end-product
ALS	amyotrophic lateral sclerosis
Ar	aryl
Arg	arginine
ARS	Alizarin Red S
Aβ	amyloid β
CD	circular dichroism
CD#	cluster of differentiation # protein
$CDCl_3$	deuterated chloroform
CEL	*N*-ε-(carboxyethyl)lysine
CJD	Creutzfeldt-Jakob Disease
CML	*N*-ε-(carboxymethyl)lysine
Da	dalton
DAF	decay accelerating factor
DCM	dichloromethane
DM	diabetes mellitus
DM-T1	type 1 diabetes mellitus
DM-T2	type 2 diabetes mellitus
DMAP	4-dimethylaminopyridine
DMF	dimethylformamide
DMPTB	*N*-phenacyl-4,5-dimethylthiazolium bromide
DNA	deoxyribonucleic acid
D_2O	deuterium oxide
δ	chemical shift
Et	ethyl
Et_2O	diethyl ether
EtOH	ethanol
ER	endoplasmic reticulum
FTO	fluorine doped tin oxide
g	gram
GC-MS	gas chromatography-mass spectrometry

GOLD	glyoxal lysine dimer
HMGB1	high mobility group box 1 protein
HPLC	high-performance liquid chromatography
HSA	human serum albumin
Hz	hertz
h*ν*	photon
IHC	immunohistochemistry
Lys	lysine
M	molar
Me	methyl
MeOD / d_3-MeOD	deuterated methanol
MeOH	methanol
MHC	major histocompatability complex
MHz	megahertz
MIF	macrophage migration inhibitory factor
mol	mole
MOLD	methylglyoxal lysine dimer
mmol	millimole
MS	multiple sclerosis
n-BuLi	normal butyl lithium
NEt_3	triethylamine
NF-κB	nuclear transcription factor κB
NFT	neurofibrillary tangle
NGFR	nerve growth factor
NMR	nuclear magnetic resonance
Nt-RAGE	*N*-terminally truncated receptor for advanced glycation end products
OSM	ovine submaxillary mucin
p	para
p-TsCl	*para*-toluenesulfonyl chloride
PAGE	polyacrylamide gel electrophoresis
PDB	protein data bank
PET	photoinduced electron transfer
PGI2	prostacyclin

ppm	parts per million
PSM	porcine submaxillary mucin
PTB	*N*-phenacylthiazolium bromide
Pyr	pyridine
r.t.	room temperature
RAGE	receptor for advanced glycation end products
RCT	randomized control trial
RER	rough endoplasmic reticulum
ROS	reactive oxygen species
SCE	saturated calomel electrode
SER	smooth endoplasmic reticulum
Ser	serine
sRAGE	soluble receptor for advanced glycation end products
tBME	*tert*-butyl methyl ether
TCR	T-cell receptor
THF	tetrahydrofuran
Thr	threonine
TMS	trimethylsilyl
TMSOTf	trimethylsilyl trifluoromethanesulfonate
Tris	tris(hydroxymethyl)aminomethane
UV	ultraviolet
ZP	zona pellucida

Contents

Chapter 1: Introduction

1.1 Glycoproteins

1.1.1 Biosynthesis and structure of glycoproteins and glycans

A glycoprotein is a protein, which has been modified to contain an oligosaccharide (also termed a glycan) covalently bonded to specific amino acid residues.[1] They are ubiquitous within the human body and serve a number of important biological functions. As such, the structure of glycans and the linkages formed between proteins and glycans have been the subject of a number of studies.[2]

Saccharides are attached to proteins in the enzymatic process called glycosylation. This is the most common example of a post-translational modification[3] and, in many cases, is vital to the proper function of the biomolecule and the organism.[3] There are two main types of glycosylation: *N*-linked glycosylation and *O*-linked glycosylation. *C*-linked glycosylation is less common and involves the addition of mannose to a tryptophan residue; it is unique in that it involves attachment to a carbon atom as opposed to the more reactive nitrogen and oxygen.[4]

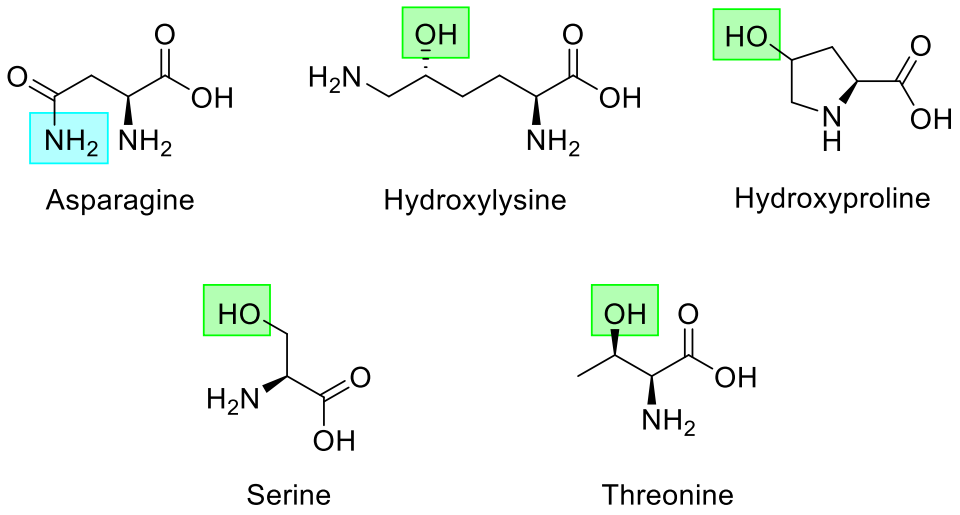
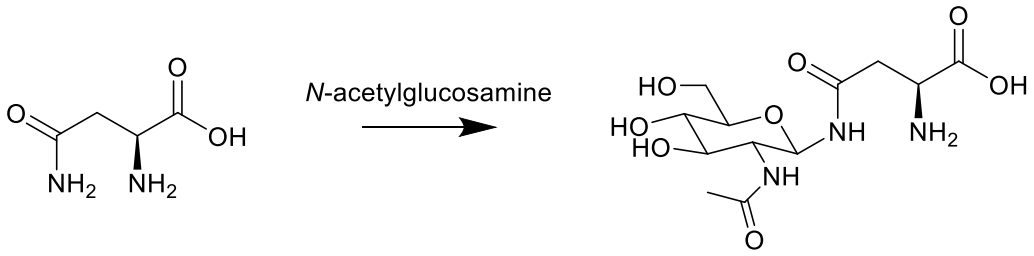


Figure 1: Common sites for *N*-glycosylation (blue) and *O*-glycosylation (green) for the amino acid residues: asparagine, hydroxylysine, hydroxyproline, serine, and threonine.[1]

In *N*-glycosylation, the saccharide is attached to the amide nitrogen atom in an asparagine residue within the protein. A very common example of glycosylation involves the binding of *N*-acetylglucosamine and the amide side chain on asparagine (Scheme 1).[1] There is only one reported example where an *N*-linked glycan is not linked to asparagine and, in this case, is instead linked to the guanidinium nitrogen on an arginine residue.



Scheme 1: Formation of a glycosidic bond between asparagine and *N*-acetylglucosamine, a very common linkage formed by glycosylation.[1]

The biosynthesis of *N*-linked glycans can be described in three phases: formation of the oligosaccharide precursor; transfer of the entire precursor to a protein; and performing alterations to the oligosaccharide chain.[3, 5, 6]

Initially, the precursor is assembled on a dolichol lipid (Figure 2), the hydrophobic portion of which is contained within the lipid bilayer of the endoplasmic reticulum (ER). The ER is an organelle found within eukaryotic cells; it is comprised of an irregular network of spaces surrounded by a membrane. It is continuous with the nuclear envelope but separate from the cytosol (the fluid within the cell). The function of the ER varies depending on cell function and type. There are two main types of ER, rough (RER) and smooth (SER). In RER the membrane is studded with ribosomes, which perform protein synthesis; SER does not have ribosomes and is primarily involved in the synthesis of lipids, phospholipids and steroids.

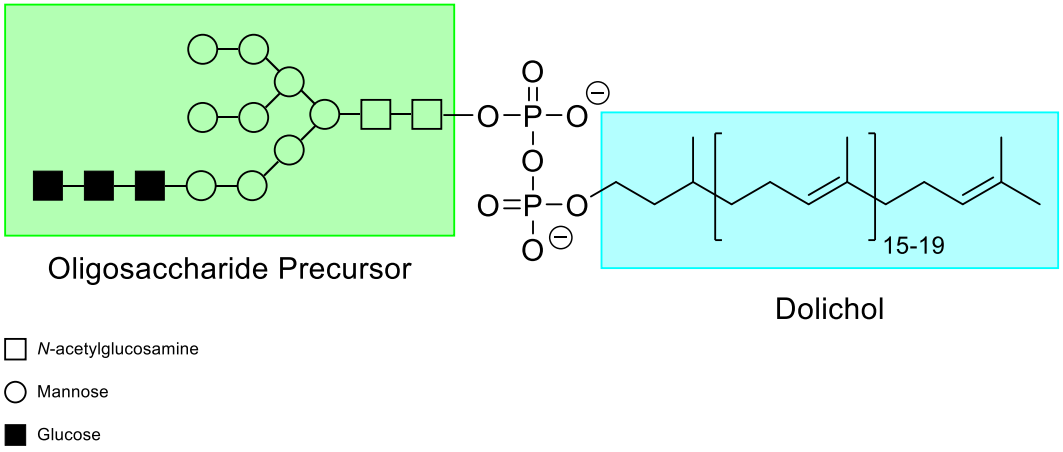


Figure 2: Dolichol lipid attached to the oligosaccharide precursor for *N*-linked glycans with a diphosphate linker. Once the precursor has been fully synthesised, the entire oligosaccharide is transferred to the target protein before it is refined as required.

With the precursor assembled, it is transferred to the target protein within the ER during or shortly after protein synthesis. The enzyme which facilitates this transfer to the asparagine residue targets a specific sequence of amino acid residues. Due to this, the *N*-linked glycans are most often well-separated from each other.[7]

Following this transfer, the protein is taken by a transport vesicle to the Golgi apparatus. The Golgi apparatus is another organelle comprised of stacks of flattened sacs, which are connected with a membranous-network. Unlike the ER, it is not connected to the nuclear envelope. Here, the sugar chain is then modified as necessary before secretion or delivery by transport vesicle to the cell membrane (Figure 3). Modifications to the glycan following attachment to a protein lead to a wide diversity of possible structures, as many as one thousand may be present in mammalian glycoproteins.[5]

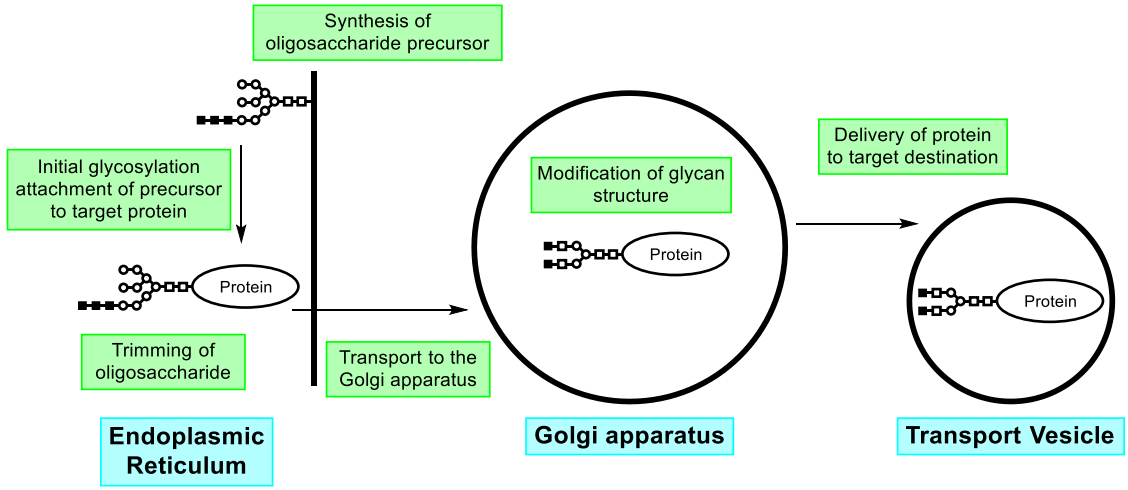


Figure 3: Schematic of the biosynthetic pathway for *N*-linked glycans and their attachment to proteins. The initial addition the oligosaccharide precursor occurs within the ER shortly after, or during, protein synthesis. Following some initial trimming of the glycan, the protein is transferred to the Golgi apparatus where further modifications take place, before the protein is secreted or transported to the cell wall.[5]

O-glycosylation is the process wherein the saccharide is attached to an oxygen atom of a hydroxyl group within an amino acid residue. It occurs commonly on the side chains of serine and threonine,[3, 5] while special types involve hydroxylysine and hydroxyproline as shown in Figure 1.[7]

There are two classifications of glycoprotein which undergo a large amount of *O*-glycosylation, these are mucins – proteins which serve as lubrication and protection from microorganisms – and proteoglycans – which play structural roles.[5] Biosynthesis of *O*-linked glycans differs from that of *N*-linked glycans as it does not include a dolichol linked precursor. While glycans found on mucins and proteoglycans vary from each other, they are both synthesised directly onto the target protein. This synthesis is initiated with the addition of a single monosaccharide; the oligosaccharide is then built up with sequential additions of further monosaccharides. The synthesis of *O*-linked glycans occurs in the Golgi apparatus and, unlike *N*-linked glycans, this attachment is performed once the protein has been assembled and folded.[7] Besides mucins and proteoglycans, some cell-surface and soluble proteins contain *O*-linked glycans.[5]

Synthesis of mucin type *O*-linked glycans starts from the addition of a single *N*-acetylglucosamine unit to a serine or threonine residue within the protein. Following the addition of this first saccharide, the chain undergoes elongation prior to termination, as with *N*-linked glycans, a large number of structures are possible.[7] However, 8 core structures have been discovered which act as a basis for further elongation (Figure 4).[7] Typically, these types of glycan are short chains and consist of only a few saccharide units and are generally clustered together on the protein.

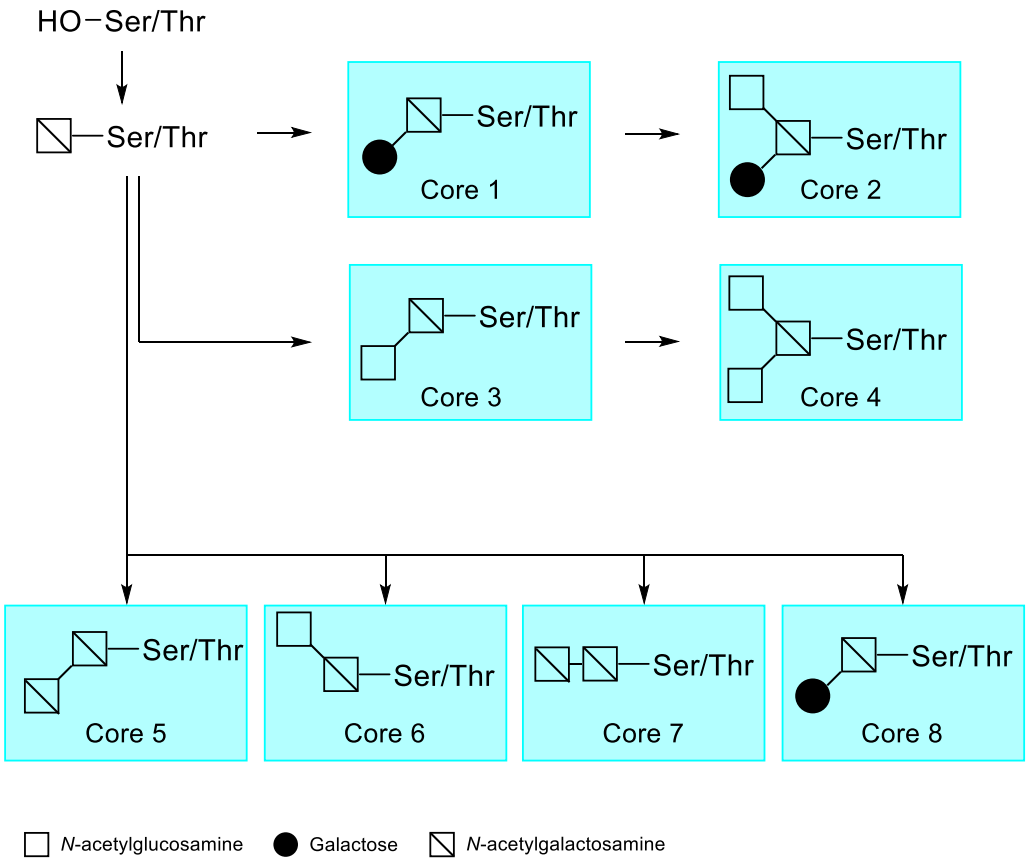


Figure 4: The 8 core structures involved in the biosynthesis of mucin type *O*-linked glycans. In these pictorials a horizontal line between two monosaccharide units represents a 1-4 glycosidic bond, an upward diagonal line represents a 1-6 glycosidic bond, and a downward diagonal line represents a 1-3 glycosidic bond. These 8 structures are elongated with the monosaccharides referenced in the figure along with fructose and mannose. Once elongation is complete the chains are terminated with sialic acid residues.[3, 7]

Non-mucin type *O*-linked glycans, such as those found on proteoglycans, are vastly different in structure. Firstly, they form long chains, with as many as 100 monosaccharide units, and secondly, these chains are linear unlike the branching structure of the mucin type glycans.[5] These linear chains are made up of repeating units of monosaccharide pairs, these glycans are known as glycosaminoglycans and their names represent the tissues in which they were initially found. Unlike mucin type glycans, these glycans are typically attached to serine or threonine residues with a xylose unit.

In some proteins, fucose and mannose can also be found linked to the hydroxyl groups of serine and threonine. Fucose may be extended to either a tetrasaccharide or disaccharide; unlike other *O*-glycosylations, the enzymes which catalyse this are found within the ER rather than the Golgi apparatus. There are also a few examples which show the *O*-glycosylation of tyrosine.[5] Another common type of *O*-glycosylation found in collagens and other proteins with a similar structure, is the attachment of a sucrose unit to hydroxylysine.

1.1.2 Effects of glycosylation on glycoproteins

Glycans play a number of roles when attached to proteins (Table 1), these include: ensuring proper protein folding; maintaining the solubility and conformation of a protein; regulating cell adhesion; binding with other proteins; and affecting embryonic development and differentiation.[3, 5, 8, 9] These effects have been studied using a range of techniques: using enzymes to remove the glycans, although this is not always possible as the sugars may be inaccessible; chemical treatment to remove glycans, however this can result in the loss of biological activity; or using cells which lack the enzymes necessary for glycosylation to synthesis the proteins to be studied.[5]

Table 1: An overview of the effects of both *N*-linked and *O*-linked glycans in both human and animal glycoproteins.[3, 5, 7, 10, 11]

***N*-linked Glycans**	***O*-linked Glycans**
• Protect proteins from digestion by proteases. • Prevent protein aggregation. • Necessary to interact with chaperones and enzymes during protein folding. • Provide rigidity to protein structure. • Orientate binding domains of proteins at cell surfaces.	• Protect proteins from digestion by proteases. • Prevent protein aggregation. • Act as receptors for recognition events. • Improve enzyme catalytic activity. • Provide rigidity to protein structure. • Supports protein structure. • Essential for protein function. • Improves solubility of proteins. • Protect against freezing of tissue. • Provide heat resistance to proteins.

Glycosylation can be an important step in protein synthesis, with glycans acting to facilitate the processes of folding or assembly. *N*-linked glycans commonly play these roles as a result of *N*-glycosylation occurring within the ER during protein translation as mentioned earlier.[3, 5, 11, 12]

For example, folding of the T cell receptor (TCR) and CD3, along with the assembly of major histocompatibility complex (MHC) (Figure 5), require multiple interactions with chaperones and enzymes, which are mediated by coordination to *N*-linked glycans.[11, 13] Two such chaperone proteins are calnexin and calreticulin; they are examples of lectins, proteins that bind to oligosaccharide moieties with high specificity. MHC class I is comprised of three sections: the heavy chain, a transmembrane glycoprotein; β_2-microglobulin, a small non-glycosylated molecule; and a peptide chain, necessary to mediate transport to the cell membrane. During assembly, calnexin binds with an *N*-linked glycan on the heavy chain in the ER. Once the heavy chain is coupled with β_2-microglobulin, it is released from calnexin and binds to calreticulin by the same *N*-glycan. This is necessary for the attachment of the peptide chain to MHC class I.[13] TCR and CD3 also bind to calnexin and calreticulin through *N*-linked glycan residues, this enables the folding of the proteins.[11]

Figure 5: The structure of the extracellular component of human MHC molecule HLA-B27[14] (PDB ID: 1HSA, image generated by Jmol version 12.2.15) showing the β_2-microglobulin chain (bottom left) and the heavy chain (right and top) which consists of four regions α1-4.

In addition to roles as receptors during protein assembly and folding, N-linked glycans can affect the structure by increasing the protein stability. In has been found that glycoproteins generally have greater stability than their non-glycosylated analogues; investigation has determined this to be a result of entropic as opposed to enthalpic effects.[12] A study of ovomucoid showed that the folded protein was stabilised with respect to the unfolded form as the *N*-glycans reduced disorder of the unfolded protein.[12]

Both *N*-glycosylation and *O*-glycosylation have been shown to effectively shield proteins from digestion by proteases. Studies have shown that the CD1 protein of mice has five sites available for *N*-glycosylation. Glycans linked at these sites act to protect the peptide chain from degredation by proteases.[13] *O*-linked glycans have been shown to hold the same role on decay accelerating factor (DAF) and CD8.[7, 11]

Glycans are important regulators of protein-protein interactions. Glycosylation can prevent aggregation of proteins on the cell surface. This is an effect of both *N*- and *O*-glycans, for example, *N*-glycosylation of CD1 prevents aggregation and ensures the correct spacing of CD1.[13] In another example, the size of sugars present on TCRs prevent the formation of TCR aggregates on membranes.[11]

As mentioned previously, mucins are a class of glycoproteins that possess regions with large amounts of *O*-linked glycosylation and are mainly produced by mucosal and glandular cells. Many mucins act to retain moisture and form a protective barrier at exposed surfaces that are not sealed by skin.[3, 5] Clusters of *O*-linked glycans allow mucins to strongly bind large amounts of water, raising the energy barrier for evaporation.[5] However, this is not the only purpose of glycans found within mucins. *O*-linked glycans are also linked to the specific flow properties exhibited by mucins, resulting in the high viscosity of fluids with high mucin concentrations. This property is essential for the lubrication of the eyes, and nasal and genital tracts by mucins.[7]

Outside of humans, *O*-glycosylation of mucins plays an important role in supporting the structure of ovine submaxillary mucin (OSM) and porcine submaxillary mucin (PSM).[7] Removing the *O*-linked glycans of OSM has been shown to have a dramatic effect on the structure, causing the extended linear structure to collapse into a globular form.[10] Another important effect of glycosylation is the increase in solubility of OSM in water. This arises as a result of glycans shielding the peptide core of the protein, which is relatively insoluble.[10] Proteins similar to mucins found in cobra venom act to improve the solubility of other proteins

within the mixture.[7] The water retention ability of mucin type *O*-glycans leads them to play an important in “anti-freeze” glycoproteins, which are found in species of Antarctic fish. These proteins have been shown to prevent the formation of ice crystals at temperatures below the freezing point of water.[3]

O-linked glycans also support the structure of nerve growth factor (NGFR). The structure of NGFR consists of a “stalk region”, a ligand binding site, and a transmembrane helix. The four identical sections of the binding domain are connected by the “stalk region” to the transmembrane helix. A single *O*-linked glycan is present on the “stalk region”; however, removal of this glycan significantly reduces the strength of ligand binding. This is suggested to be a result of the glycan directing the binding site away from the cell surface.[7]

Cell surface glycans have been shown to have important roles in recognition and cell adhesion. This is particularly evident in fertilization where *O*-glycans facilitate the binding of the sperm to the egg.[3, 7] The egg is surrounded by a glycoprotein coat, the zona pellucida , which is made up of four primary glycoproteins, called ZP1, 2, 3, and 4. Removal of N-glycans from the zona pellucida has no effect on the binding of spermatozoon and oocyte, whereas elimination of *O*-glycans prevents the binding of a sperm cell.[3] It has been shown that, in mice, ZP3 acts as the receptor for spermatozoa and that *O*-linked glycans are essential for this recognition.[7] The glycosylation pattern and structure of glycans of ZP3 has been observed to vary between different species, suggesting that *O*-glycans play a role in the species specific sperm-egg binding.[3, 7]

Glycans also play an important role in protein recognition within the immune system. Recognition of foreign proteins and the subsequent initiation of immune response is a basic action of the immune system. An epitope is the domain of an antigen that allows for immunological recognition by antibodies and immune systems cells. Glycosylation patterns present on epitopes have a strong effect on the recognition by antibodies.[7] Glycans are also implicated in many of the steps required for recognition of antigen presenting cells by T cells.[11]

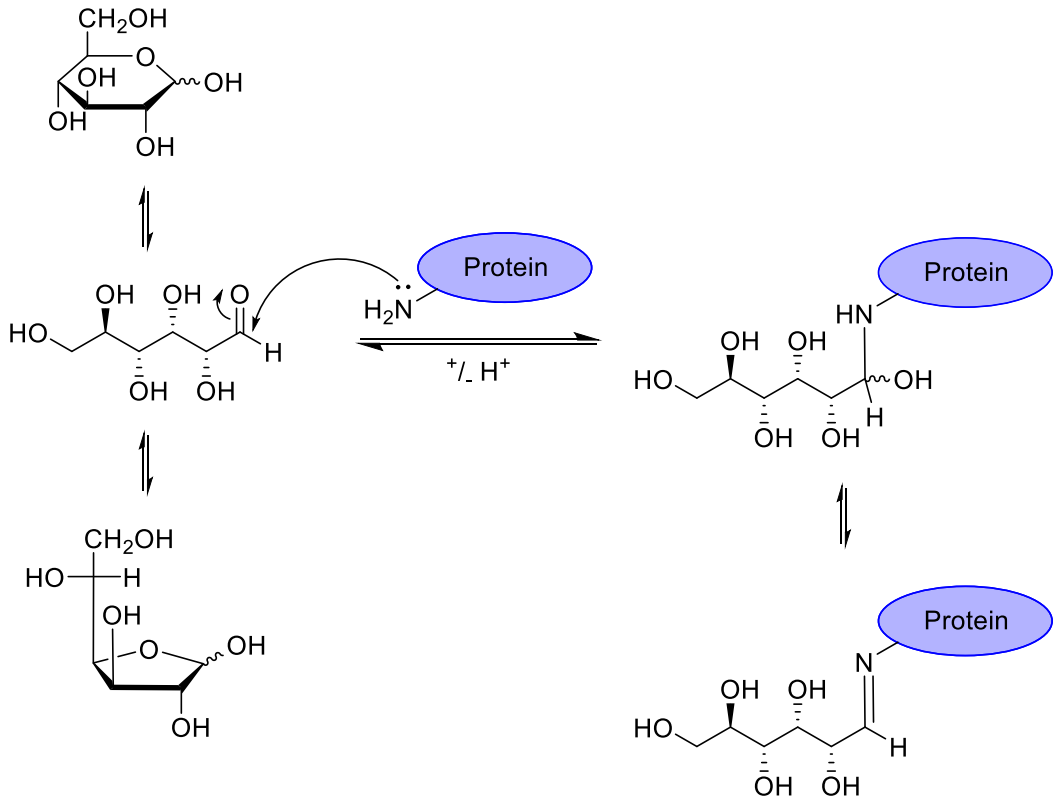
Cell adhesion molecules CD2 and CD48 in rats, or CD58 in humans, facilitate the binding of T cells with antigen presenting cells. It is proposed that CD2 and CD48 or CD58, present on the cell surface act to align T cells with the target cell so that TCRs can bind with antigens indentified by MHC molecules.[13] Whilst *N*-linked glycans present on CD2 and CD48 are not located close to the binding domains of the proteins, they promote the orientation of these binding regions away from the cell surface and prevent protein-protein interaction between

CD2 and CD48 on either the same, or other, cells. This is necessary for CD2 and CD48 to effectively adhere to antigen presenting cells.[13]

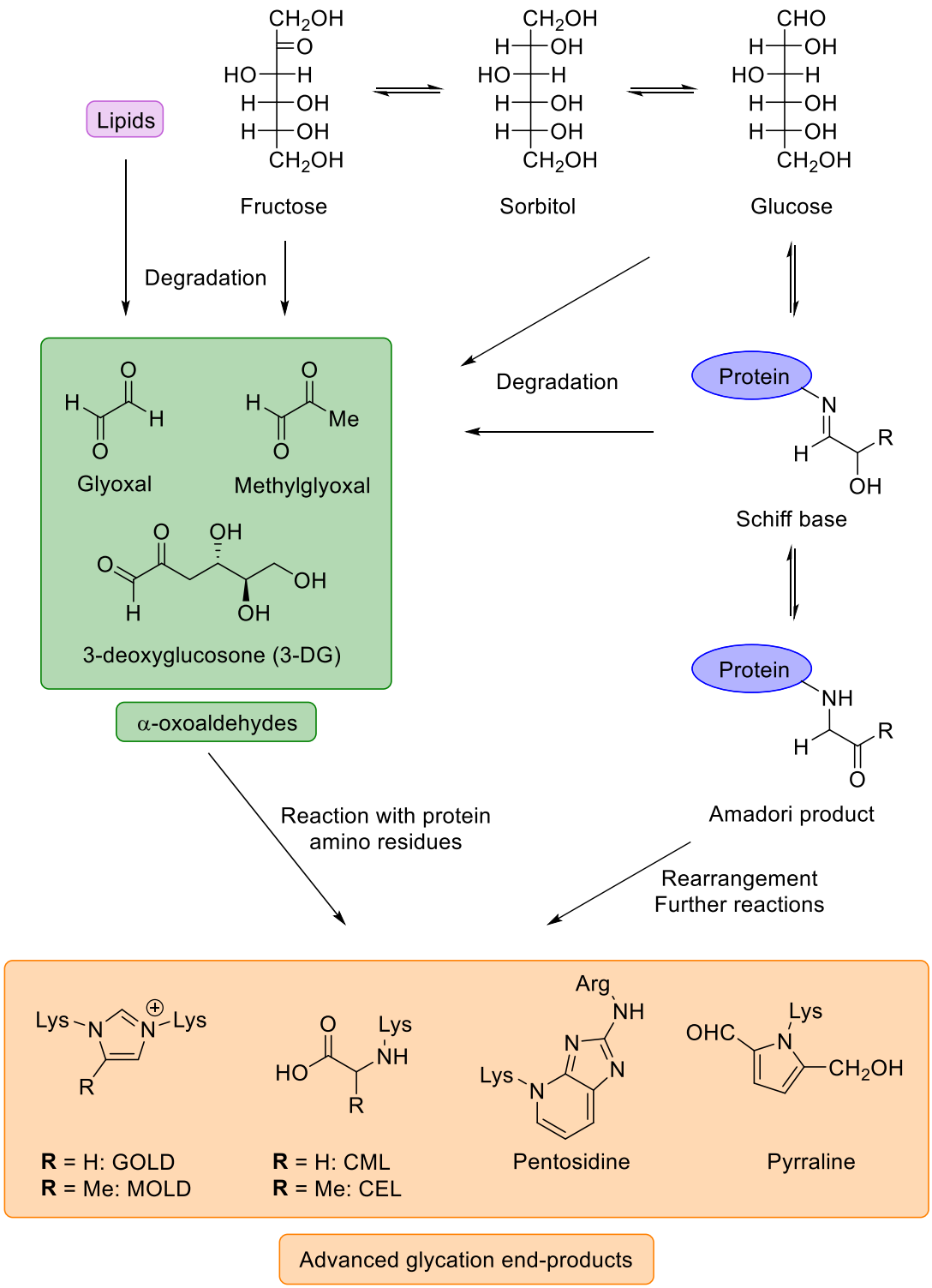
These examples have highlighted the diverse, and often essential, roles of glycoproteins and how glycans affect the structures and facilitate proper function of the biomolecules they are attached to.

1.1.3 Glycation and AGEs

As discussed above, glycosylation is an important post-translational modification of proteins and greatly affects the function and structure of the protein. Glycation is the non-enzymatic equivalent of glycosylation and is the process in which reducing sugars – primarily fructose, glucose and galactose[15-17] – react directly with proteins resulting in the covalent attachment of the saccharide. Such reactions occur slowly *in vivo* and as such are mostly observed when blood sugar levels are raised, for example in patients with diabetes mellitus. In non-diabetic subjects glycation by sugars can still be observed, however it is mostly present on long lived proteins.[5]



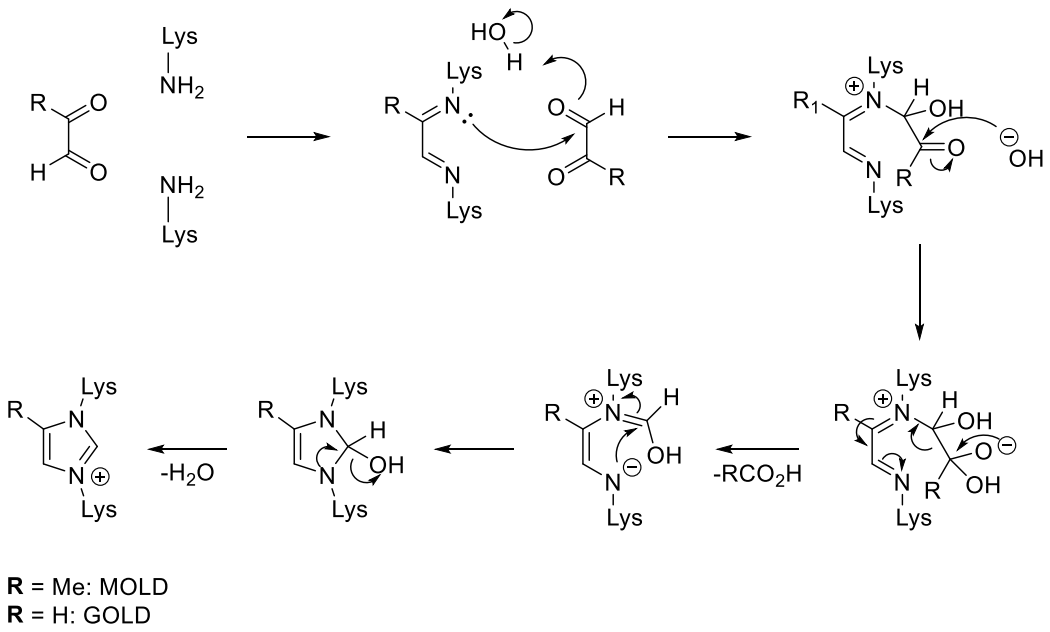
Scheme 2: The linear form of D-glucose, in equilibrium with the cyclic forms, reacts with an amino group of an amino acid residue on the peptide chain of a protein. This non-enzymatic reaction is termed glycation, it occurs slowly *in vivo* and the products are only commonly found in those with raised blood sugar levels.



Scheme 3: Two possible pathways for the formation of advanced glycation end-products from D-glucose and D-fructose, which can be interchanged via the polyol pathway. Direct degradation of glucose or fructose results in the formation various α-oxoaldehydes, which lead to the formation of the AGEs shown and protein-protein cross-linkages.[18, 19] These α-oxoaldehydes can also form as the result of lipid peroxidation.[20, 21] The rearrangements and further reactions that follow the coupling of glucose with amino residues on proteins provide another route to these AGEs among others.[15] Abbreviations used in the scheme are: GOLD (glyoxal lysine dimer), MOLD (methyl glyoxal lysine dimer), CML (*N*-ε-(carboxymethyl)lysine), and CEL (*N*-ε-(carboxyethyl)lysine).

This reaction involves acyclic sugars and terminal amino groups of amino acid residues in the peptide chain of the protein. The open-chain form of sugars exists in equilibrium with the pyranose and furanose forms in solution. In this linear form the carbonyl is free and able to react with an amino group (Scheme 2). As glycation is non-enzymatic, there is no control over the reaction sites. Studies have shown that non-specific addition of sugars is a major contributor to the damage of biomolecules and can impair or destroy the functionality and cause deviation from the correct structure by encouraging misfolding.[22, 23]

The Schiff base formed by glycation is termed an early glycation product, further rearrangements and reactions lead to the formation of advanced glycation end-products (AGEs) (Scheme 3). The addition of D-glucose via glycation acts as the first step along one pathway toward AGEs. After the Schiff base has undergone an Amadori rearrangement, further reactions and rearrangements ultimately yield AGEs;[15] this set of reactions is called the Maillard reaction and is one of the causes of browning observed on heating meats and other foodstuffs.[15]



Scheme 4: Mechanism for the formation of MOLD and GOLD, involving the reaction of methylglyoxal or glyoxal with two lysine residues respectively.[18]

Glucose and fructose can be interchanged within the body via the polyol pathway. This conversion is mediated by two enzymes: aldose reductase (to reduce glucose to sorbitol); and sorbitol dehydrogenase (to oxidise sorbitol to fructose). However, as aldose reductase has only a weak affinity for glucose, this conversion only occurs when the concentration of glucose in the blood is particularly high.[8] Glucose, fructose or the Schiff base can degrade, leading to the

formation of α-oxoaldehydes (or α-dicarbonyls) further reactions then lead to the formation of AGEs. Fructose reacts with proteins much more rapidly than glucose to form AGEs.[15, 24] Peroxidation of lipids can also lead to the formation of α-oxoaldehydes providing another route to AGE precursors.[21] Scheme 3 shows *N*-ε-(carboxymethyl)lysine (CML) and *N*-ε-(carboxyethyl)lysine (CEL), which can be formed when the ε-amino group of lysine reacts with glyoxal or methylglyoxal respectively.[15] Glyoxal and methylglyoxal can also react with two lysine residues to form glyoxal lysine dimer (GOLD) or methylglyoxal lysine dimer (MOLD) (Scheme 4).[18] GOLD and MOLD can act as non-specific protein cross-links – another example of AGEs – resulting in the aggregation of proteins. Such cross-links can also be formed following protein glycation by glucose.[15]

1.1.4 AGEs and their biological impact

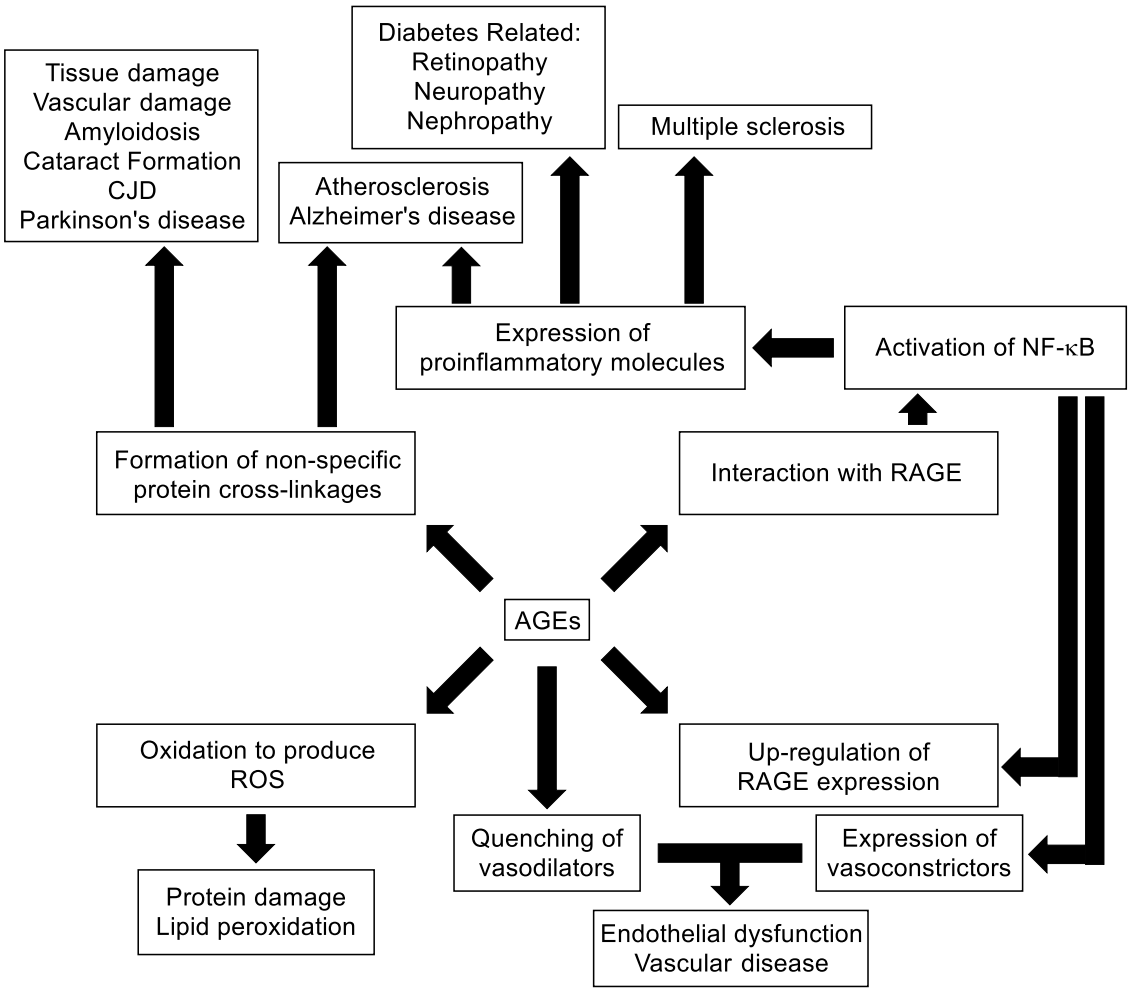
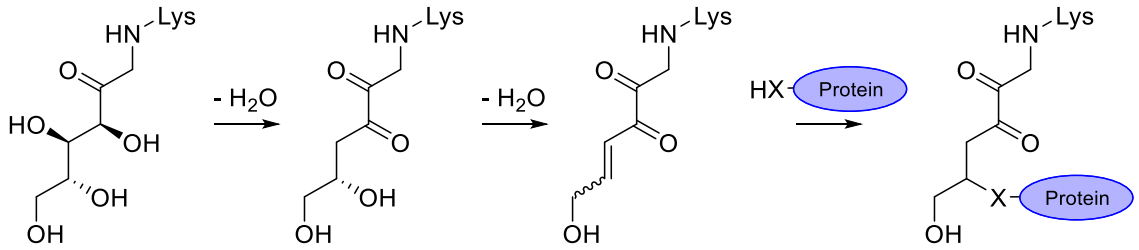


Figure 6: A flow chart depicting some of the possible pathological pathways initiated by the formation of AGEs and resulting interactions. AGEs possess both direct toxic effects, by reacting with vasodilators, degradation to reactive oxygen species resulting in reactions with proteins and lipids, and formation of non-specific protein cross-links. Protein cross-linkages have been implicated in a range of neurodegenerative disease and lens disorders. Interaction of AGEs with the receptor for AGE leads to activation of inflammatory pathways, which participates in the pathogenesis of chronic inflammatory diseases. This interaction has also been shown to have an involvement in the pathogenesis and development of diabetes related diseases.[15, 22, 24-29]

The presence of AGEs and their interactions have been implicated in the pathogenesis and progression of both diabetes and age related diseases such as: atherosclerosis, Alzheimer's disease, rheumatoid arthritis, chronic kidney disease, vascular disease, multiple sclerosis, neuropathy, nephropathy, retinopathy, and others.[15, 22, 24, 27-29]

As previously stated, AGE levels in tissue and blood, have been linked to hyperglycaemia, which may arise as a result of poorly controlled diabetes. This can be shown by the significantly increased tissue and blood AGE concentrations in diabetic patients with respect to control subjects. Amongst diabetics, higher AGE concentrations are observed in those who have developed complications, such as microvascular disease. However, it has also been shown that AGE levels rise in healthy subjects with age and are present even in young children. Studies have shown that older non-diabetics with no health complications have a greater accumulation of AGEs in tissues than their younger diabetic counterparts.[15] From this it has been suggested that it is not the concentration of AGEs that leads to the development of disease, but the rate of accumulation;[30] although, AGE concentration has been shown to have a number of detrimental effects at the cellular level (Figure 6).[15, 24]



Scheme 5: Formation of a protein cross-linkage by dehydration of the Amadori product resulting from the glycation of a protein by D-glucose.[31]

Commonly, formation of AGEs leads to the covalent cross-linking of proteins. This mostly affects stable proteins with long lifetimes, such as collagens. As shown in Scheme 4, α-oxoaldehydes can react with lysine residues on separate proteins to form covalent links between two proteins via the AGEs: MOLD and GOLD.[19] Another important mechanism for the formation of cross-links goes via the Amadori product formed after glycation. Successive dehydrations of the Amadori product produces a 1,4,5-trideoxy-1-alkylamino-2,3-hexulos-4-ene, which can then go on to react with the amino group side chains of lysine or histidine or the thiol side chain of cysteine to form protein cross-links.[31] Cross-linking of proteins can occur naturally in the body by the enzymatic action of lysyl oxidase, although there is no evidence that such cross-links are included in the increased cross-linking observed in diabetic subjects.[15] The protease-resistance of these cross-links leads to accumulation of proteins and amyloidosis,

the process by which soluble proteins are rendered insoluble by structural alterations and conglomerate in extracellular space of tissues.[15, 24]

Formation of non-specific protein cross-links by AGEs result in the stiffening of tissue by the accumulation of matrix proteins.[15] Post-mortem histological studies on human aortas have shown there is a correlation between the stiffness of the aorta and accumulation of AGEs within the tissue. Other results of cross-link formation include the thickening of basement membranes in capillaries, and development of atherosclerosis.[15]

AGEs have been shown to interact with the endothelium, the surface of blood vessels, by binding with a number of cell surface proteins. Amongst these receptors are: lactoferrin, oligosaccharyl transferase 48 (AGE-R1), 80K-H protein (AGE-R2), galectin-3 (AGE-R3) and the most well characterised, the receptor for AGE (RAGE) along with others.[15, 28]

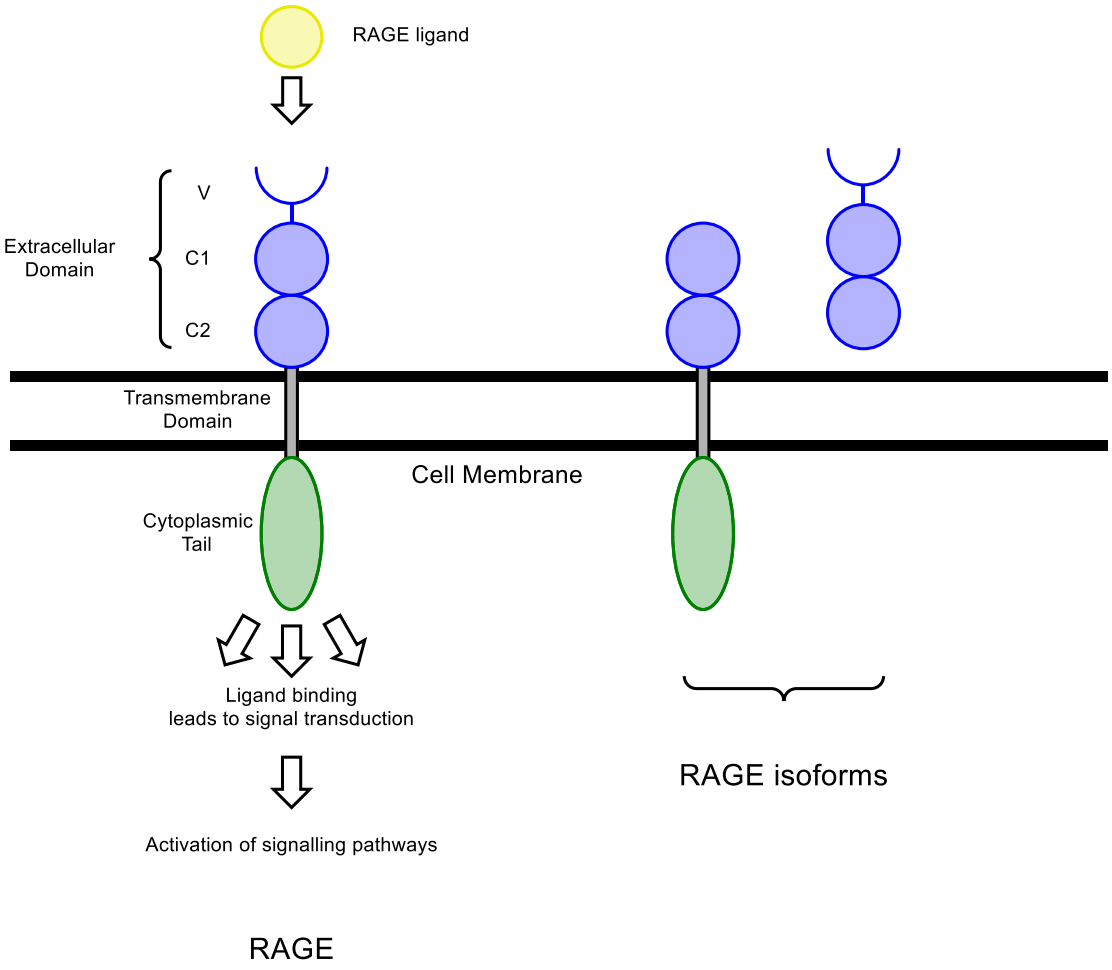


Figure 7: A schematic diagram of the structure of RAGE and activation by a RAGE ligand. Two of the main isoforms of RAGE are shown on the right side of the image: *N*-terminally truncated RAGE (Nt-RAGE) (left); and soluble RAGE (sRAGE) (right). These isoforms may be generated by alternative splicing or, in the case of sRAGE, protease cleavage from the cell surface.

RAGE is a member of the immunoglobulin superfamily, a group of proteins that share structural features with antibodies (immunoglobulins). The full 404 amino acid long RAGE protein, consists of three primary regions: the extracellular domain, which in turn possess three immunoglobulin-like moieties termed the V domain and the C1 and C2 domains; the transmembrane region, made up from a single hydrophobic chain; and a 43 amino acid cytoplasmic tail (Figure 7).[25, 26, 32] The V domain is necessary for the binding of ligands by RAGE and the cyctoplasmic tail is used for signal transduction.[26] Isoforms of RAGE exists, the most common of these include, a soluble version which lacks both the transmembrane and cytosolic domain, named soluble RAGE, and an *N*-terminally truncated form, which lacks the V immunoglobulin domain. Both can be synthesised by alternative splicing of RAGE; however, sRAGE may also be formed by cleavage of RAGE from the cell surface by proteases.[25] A number of polymorphs of RAGE, where an amino acid has been substituted for another, have been identified and linked to increased risk and onset of certain diseases.[25]

It has been shown that RAGE interacts with a number of ligands besides AGE; this occurs by recognition of the three dimensional structure rather than an amino acid sequence, identifying RAGE as a pattern-recognising receptor.[25, 26] These ligands include: amyloid beta (Aβ), a biomarker of Alzheimer's disease (AD); amyloid fibrils, formed by the deposition of proteins in amyloidosis; S100 proteins, which are calcium ion binding proteins that play a number of roles in the cell cycle; and high mobility group box 1 protein (HMGB1), which is involved in organisation of chromosome architecture and is free to move between the nucleus and cytoplasm. Binding of these ligands, or with AGEs, leads to activation of RAGE and a number of cell signalling pathways (Figure 7).[26]

An important signalling pathway involves RAGE mediated activation of nuclear transcription factor κB (NF-κB). NF-κB is present in the cytoplasm bound to the inhibitor molecule IκBα; activation of RAGE leads to phosphorylation and degradation of IκBα, causing the release of NF-κB, which is then translocated into the nucleus.[26] This results in the production of proinflammatory cytokines (molecules used in cell signalling), reactive oxygen species, IκBα, and increased expression of RAGE on the cell surface.[24-26] This pathway is evoked as part of immune responses; RAGE and other cell surface receptors bind to 'danger signals', which leads to the activation of infection-associated inflammation.[25] However, when this pathway is activated by ligand interaction with RAGE, the feedback loop that regulates NF-κB activation can become overwhelmed as production of new IκBα is insufficient to contain NF-κB within the cytoplasm.[26] As activation of NF-κB increases expression of RAGE, the number of sites for the

binding of RAGE ligands is also raised and leads to further activation of NF-κB and amplification of the signal.[33] This prevents inhibition of NF-κB and leads to the continued up-regulation of proinflammatory signalling molecules and sustained inflammation resulting in the development of chronic inflammatory conditions such as those that arise as complications of diabetes mellitus.[15, 25, 26, 28, 33, 34] This has been illustrated in diabetic animal models as mice deficient in RAGE, or with RAGE activation blocked, showed a significant reduction in the development of complications and partial protection from functional deficits, such as nephropathy and neuropathy, as compared to control mice and mice overexpressing RAGE.[26]

RAGE may be found on the surface of many cell types although expression is low in most healthy adult cells except those of the lung and skin, which express RAGE constantly throughout life. However, it has been found that RAGE expression is up-regulated by the presence of large quantities of RAGE ligands.[26] This up-regulation of RAGE has been observed in a variety of diseases, such as diabetes mellitus (DM) and Alzheimer's disease.[25]

Activation of RAGE also induces the production of the vasoconstrictor endothelin-1 as a result of the NF-κB pathway.[26, 28] Expression of endothelin-1 combined with AGE inhibition of vasodilators leads to a more powerful vasoconstrictive effect. There are two vasodilators that are known targets of AGEs; these are nitric oxide (NO), which is rapidly quenched by the presences of AGEs[35]; and prostacyclin (PGI2), which studies[36] have shown has decreased production in the presence of AGEs. As it has been shown that increased AGE concentration in diabetic subjects correlates with severity of endothelial dysfunction, these effects may play an important role in the vascular complications of DM.[28]

As stated above, the activation of RAGE results in the generation of reactive oxygen species (ROS), an increase in oxidative stress, and hence a reduction in antioxidants. The direct oxidation of AGEs, Amadori products, and lipids is also associated with the formation of ROS, particularly oxygen free radicals, superoxide and hydrogen peroxide.[24] Studies have highlighted that the generation of ROS is important in perpetuating the activation of NF-κB, as signal transduction and the effects of RAGE activation have been shown to be inhibited by anitoxidants.[28] ROS can also lead directly to cell damage by reacting with proteins and lipids.[24]

Much work has focused on the role of AGEs in neurodegenerative disorders as a result of aging. Early work focused mostly on Alzheimer's disease, while more recent investigations have considered the role of AGEs in multiple sclerosis (MS), and amyotrophic lateral sclerosis (ALS).[22, 24]

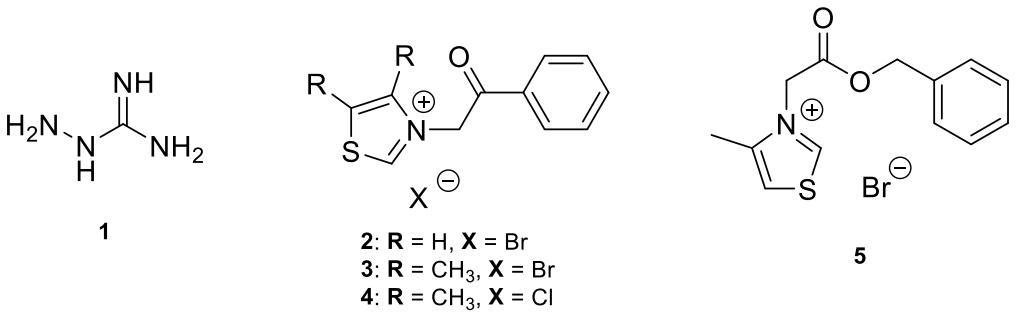
Intracelluar deposits of hyperphosphorylated tau protein, called neurofibrillary tangles (NFTs), along with aggregates of both actin and associated proteins called Hirano bodies, are found within the neurons of the AD brain. These NFTs are an important biomarker for AD.[24] Studies investigating the presence of AGEs within neurons have shown that the proportion of neurons which contain AGEs increases with aging, and, in those affected, progression of AD. Additionally, a connection between AGE accumulation and early NFT formation has been proposed following evidence that showed all neurons in which AGEs were present contained hyperphosphorylated tau.[37]

Senile plaques, extracellular deposits of Aβ oligmers, are found in the brains of those with Alzheimer's disease. It has been shown that most plaques both contain AGEs and exhibit an increased level of AGE formation.[24] Studies have shown that the formation Aβ oligmers is significantly accelerated by the formation of AGEs and AGE cross-linkages between Aβ monomers. Further investigation has suggested that AGE formation is important in the initial stages of plaque development.[24] As both AGEs and Aβ are recognised by RAGE, expression of the receptor is found throughout the AD brain and it is thought that RAGE has an important role in AD.[22]

Along with AD, AGEs have been implicated in other neurodegenerative diseases. The up-regulated expression of RAGE and its ligands has been reported in active lesions of MS while normal expression is observed in inactive ones; this was found in both patients with MS and animals models of MS and it is thought that this plays a role in the increased inflammatory response in MS.[38] In Parkinson's disease Lewy bodies, abnormal protein aggregates observed inside nerve cells, and aggregates of α-synuclein are triggers of cell death.[24] AGE cross-linking of α-synuclein has been suggested to be one of the initial steps in the formation of Lewy bodies and, therefore, the pathogenesis of Parkinson's disease.[39] Similarly, AGE cross-linked proteins, AGE alteration of proteins leading to improper folding, and RAGE expression have been observed in the brain tissue of patients with Creutzfeldt-Jakob disease and is thought to contribute to the progression of the disease.[22, 24, 40]

1.1.5 Treatments to target AGEs

As the pathogenic role of AGEs in a range of diseases has become apparent, interest has risen in investigating their potential as therapeutic targets. A number of *in vitro* animal model studies and clinical trials have been conducted to investigate treatments that specifically inhibit the formation of AGEs, remove preformed AGEs, and prevent their interactions, with the aim of mitigating the previously discussed effects (Figure 8).[26, 29, 41-46]



1

2: **R** = H, **X** = Br
3: **R** = CH_3, **X** = Br
4: **R** = CH_3, **X** = Cl

5

Figure 8: The structures of some AGE inhibitors and AGE cross-link breakers that have been subject to investigation in clinical trials, animal studies, and *in vitro* studies. These molecules are: aminoguanidine (1), PTB (2), DMPTB (3), alagebrium (4), and C36 (5).

Aminoguanidine (**1**) received great interest for its potential as an inhibitor of AGE formation. As it reacts with the carbonyl groups of sugars, it prevents these sugars from reacting with proteins hence blocking the formation of AGEs.[29] In animal models of diabetes, aminoguanidine has been shown to reduce the characteristics of nephropathy and neuropathy; however, these results were not reproducible in a study on early nerve changes in diabetic baboons.[15, 29, 47] Clinical studies of aminoguanidine have investigated its effects in double-blind randomized control trials (RCTs) of both type 1 (DM-T1) and type 2 (DM-T2) diabetes mellitus.[48, 49] The trial on patients with DM-T2 was terminated early due to safety concerns from side-effects of aminoguanidine, while the trial on patients with DM-T1 did not produce any statistically relevant results on the beneficial action of aminoguanidine. Despite the results of animal and *in vitro* studies, it now seems unlikely aminoguanidine will be used in therapeutic applications.[29, 41]

It has been shown that AGE cross-link breakers such as *N*-phenacylthiazolium bromide (PTB) (**2**), *N*-phenacyl-4,5-dimethylthiazolium bromide (DMPTB) (**3**), and alagebrium (**4**) can alleviate and reverse complications of DM. As a prototype cross-link breaker, PTB has shown the ability to selectively cleave cross-links formed by the degradation and further reaction of Amadori products (via the mechanism shown in Scheme 5).[31, 42] DMPTB and alagebrium have been developed in light of the instability of PTB. Studies have shown the ability of DMPTB to break, *in vitro*, AGE cross-links of α-crystallin – a protein found in human lenses – preformed both *in vitro* and *in vivo*, while also inhibiting the formation of further cross-links by immediate cleavage after formation.[45] This leads to the reversal of protein accumulation, an important step in cataract formation. C36 (**5**) has been shown to have AGE cross-link cleaving action *in vitro* and improvements in arterial elasticity and endothelial dysfunction in diabetic rats.[42]

Alagebrium has been the subject of a wide range of studies, which have shown the ability of treatment with alagebrium to reduce the pathological effects of AGE cross-linking in a number

of associated diseases and disorders.[41, 42] Animal studies have shown that alagebrium treatment reduced arterial stiffness, prevented the accumulation of AGE cross-linked collagens in the aorta, and ameliorated the effects of diabetic nephropathy in rats and mice with induced diabetes. Investigations have also shown the beneficial effect of alagebrium on vascular stiffness in elderly and diabetic dogs leading to improved vascular function. Similar improvements in vascular and arterial function were observed in aging, non-diabetic, rhesus monkeys. Alagebrium has also been shown to have a beneficial effect on retinopathy, lens disorders, neuropathy and erectile dysfunction in DM.[42]

In clinical studies, improvement of cardiovascular function was observed in one RCT on alagebrium treatment of hypertensive patients,[50] although the effects were not reproduced[51] in those with heart failure. Smaller, non-controlled trials have provided evidence on the beneficial effects of alagebrium treatment on arterial stiffness, endothelial and ventricular function; however, no improvement was seen in exercise performance.[41, 52]

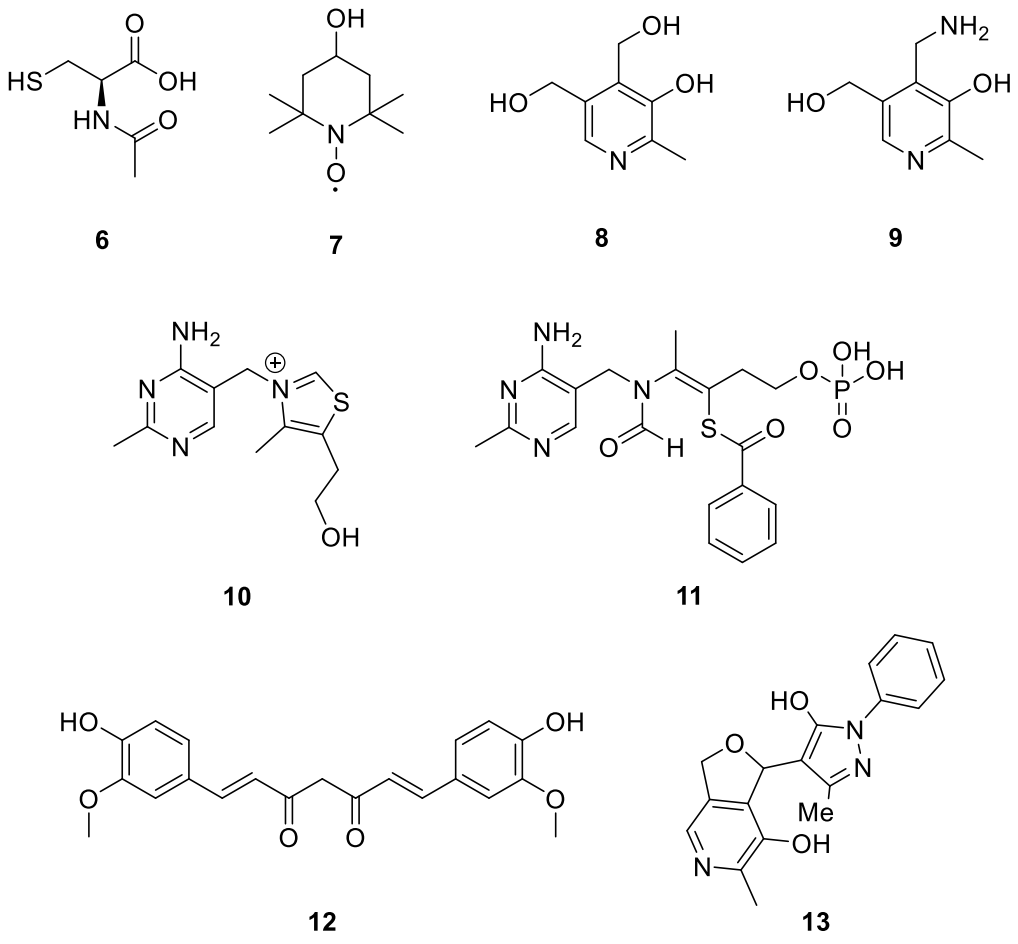


Figure 9: The structures of a number of molecules that have been investigated for their antioxidant or ROS scavenging applications as possible inhibitors for the formation of AGEs. These molecules are: *N*-acetyl cysteine (6), TEMPOL (7), pyridoxine (8), pyridoxamine (9), thiamine (10), benfotiamine (11), curcumin (12), and TM2002 (13).[41, 43]

Antioxidants can act as scavengers for ROS, reducing oxidative stress and removing possible AGE precursors, effectively inhibiting the formation of AGEs. *N*-acetyl cysteine (**6**) acts to reduce oxidative stress by removing ROS, although high levels of *N*-acetyl cysteine will induce apoptosis, while curcumin (**12**) acts as a non-toxic scavenger of superoxide ions.[43] TEMPOL (**7**) a powerful antioxidising agent has been shown to lessen the symptoms of endothelial dysfunction as a result of oxidative and endoplasmic reticulum stress.[53] Testing of TM2002 (**13**) *in vitro* has highlighted its ability to inhibit AGE formation by acting as both a scavenger of radicals and by metal chelation, preventing metal catalysis of AGE formation.[54] Additionally, *in vivo* studies showed low toxicity while also improving renal function.

Two forms of vitamin B6, pyridoxine (**8**) and pyridoxamine (**9**), vitamin B1, thiamine (**10**), and its derivative, benfotiamine (**11**), have all exhibited antioxidant activity and AGE formation inhibition *in vitro* and in animal studies.[41] Some evidence exists for the activity of pyridoxamine *in vivo* in patients with both DM-T1 and DM-T2, shown by reduced plasma levels of the AGEs, CML and CEL.[41, 55] Beneficial effects of the other antioxidants listed are yet to be established with respect to diabetic complications.[41]

Another possible target for treatments is to block the interaction of RAGE and its ligands. Soluble RAGE (sRAGE), as mentioned previously, is an isoform of RAGE that is not bound to the cell membrane.[25] sRAGE acts a decoy, binding with AGEs and other RAGE ligands to prevent their binding with the cell surface RAGE, which in turn prevents signal transduction and activation of proinflammatory pathways via NF-κB.[26]

Using sRAGE in diabetic mice and rats has shown a variety of benefits in inflammatory conditions.[26, 29, 56] It has been demonstrated that sRAGE can suppress diabetic atherosclerosis in mice by blockage of RAGE, successfully inhibiting the formation of atherosclerotic lesions, stabilising the area of pre existing lesions, and preventing expansion of scar tissue within the arteries – a feature observed in chronic atherosclerosis.[57, 58]

Other studies have shown an improvement in vascular complications though the suppression of vascular inflammation.[59] This was shown to be a result of blockading RAGE, as other antagonistic features, such as glycation as measured by the level of glycated haemoglobin, were unchanged.

Loss of pain perception in diabetic mice was shown to be reduced following treatment with sRAGE, indicating that sRAGE inhibited dysfunction of neurons and diabetes related neuropathy.[60] Delayed wound healing is another common complication of DM and can result

in amputation of digits or limbs.[56] Blockage of RAGE activation by sRAGE was shown to restore wound healing in diabetic mice.[61] Furthermore, evidence shows that sRAGE can successfully reduce the effects of nephropathy by preventing inflammation in the kidneys of mice.[44]

Studies comparing RAGE deficient mice with those treated by sRAGE showed that treatment with sRAGE conferred better protection from pathological events.[60] This result suggests that sRAGE is able to prevent the interaction between ligands and other receptors, in turn implying that RAGE is not the only receptor for AGEs and other ligands, which is important in inflammatory conditions.[26, 60]

1.2 Recognition

1.2.1 Molecular recognition and its applications

Molecular recognition is a selective interaction between two molecules, a receptor and a substrate. The selectivity of the binding event relies on the complementary relationship between the structures of the receptor and the substrate. Highly specific recognition is observed within biological systems, in the binding of biological receptors and their ligands (such as RAGE and AGEs), enzymatic catalysis, transcription of proteins, and antigen recognition by antibodies. Many synthetic receptors have also been designed for use in self-assembly, molecular sensors, and molecular machines. Designing these molecules requires the manipulation of supramolecular interactions, such as hydrogen bonding, electrostatic interactions, π-stacking, or van der Waals forces.[62]

Some common examples of molecular recognition are shown in Figure 10. Crown ethers can selectively recognise group 1 metal ions, Figure 10 shows 15-crown-5 (**14**) binding with a sodium ion and 18-crown-6 (**15**) binding to a potassium ion. The selectivity arises from ion size; the larger potassium ion cannot interact with the 5 binding sites of 15-crown-5 without buckling the ring, while a sodium ion is too small to interact with all 6 binding sites of 18-crown-6 simultaneously, thus 18-crown-6 binds more strongly to potassium than sodium and vice-versa for 15-crown-5. The difference in binding energies gives rise to the selectivity of the ethers. Alongside the crown ethers, Figure 10 shows a biological example in the interaction between the nucleobases. Complementary hydrogen-bonding motifs allow for the specific formation of guanine-cytosine pairs and adenine-thymine pairs. The pairing of guanine with cytosine (**16**) and adenine with thymine (**17**) allows for the formation of the maximum number hydrogen bonds, as such these pairs form selectively.

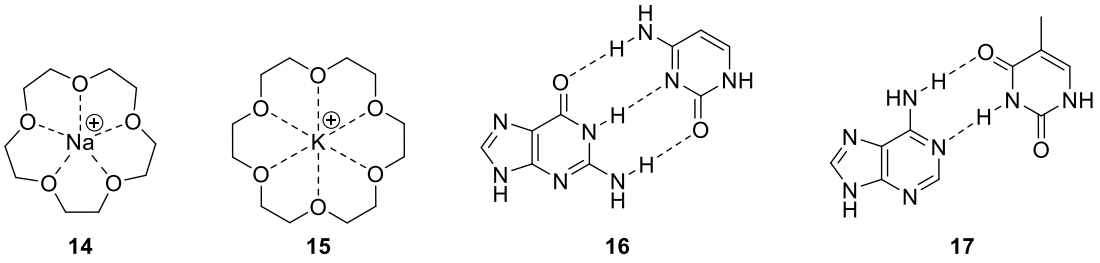


Figure 10: Common examples of molecular recognition. The crown ethers, 15-crown-5 (14) and 18-crown-6 (15), can selectively bind to metal ions by electrostatic interactions (as shown by the dashed bonds), 15-crown-5 is selective for Na^+ while 18-crown-6 is selective for K^+. This is a result of varying binding constants that arise from the size of the ions. Also shown are the DNA base pairs, guanine-cytosine (16) and adenine-thymine (17), H-bonds (as shown by dashed bonds) form between the complementary acceptors and donors to give selective formation of these pairs.

Receptors are essential components of sensors. A molecular sensor (Figure 11) is defined as a device, which transforms chemical information into an analytically useful signal.[63] The signal may come in a range of forms, for example, nuclear magnetic resonance (NMR) spectroscopy, mass spectrometry, electrochemical, or visible in the form of fluorescence, phosphorescence or a colour change. This signal is a result of the binding of a substrate by the receptor, returning a macroscopic response to the microscopic input.

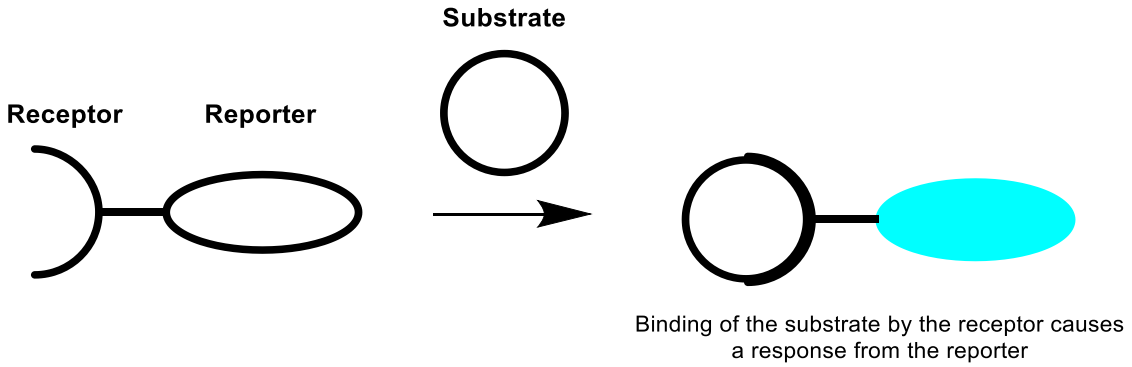
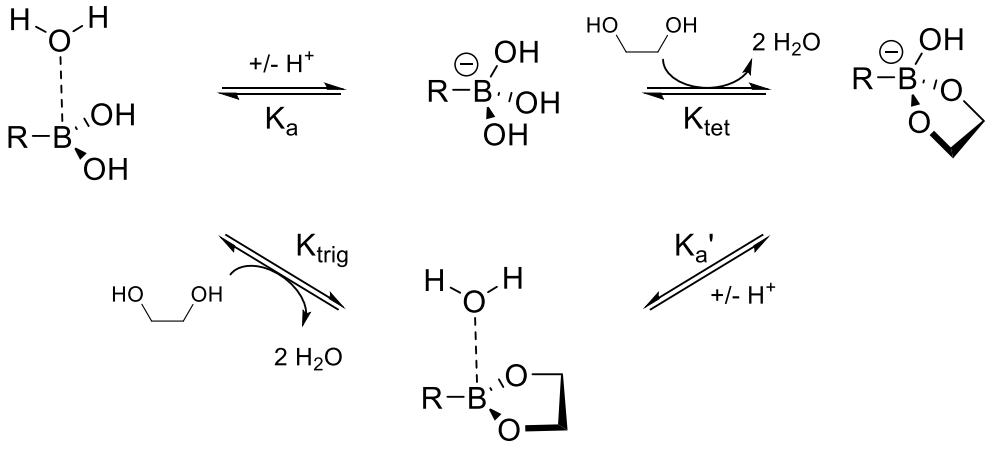


Figure 11: Principle of a molecular sensor comprised of a receptor to specifically bind with a target molecule and a reporting unit to produce a macroscopic signal.

1.2.2 Carbohydrate receptors

Owing to the biological significance of carbohydrates, the development of methods to detect the presence and concentrations of sugars has been the subject of a large body of investigation. Carbohydrate sensors have a range of applications, in both industrial and medical situations. The medical ramifications of improper sugar metabolism have already been discussed, monitoring sugar levels can be important for following the progress of diseases. In industry sugar detection can have a number of uses such as measuring the progress of fermentation.[64]



Scheme 6: The cycle of equilibrium reactions that is involved in the formation of a boronate ester from a boronic acid and *cis*-diol. The formation constants K_{trig} and K_{tet} represent the formation of a boronate ester from the trigonal planar and tetrahedral forms of the boronic acid respectively. K_a and K_a' are the acidity constants for boronic acid and boronate ester respectively.

Many carbohydrate sensors have utilised the ability of boronic acids to bind reversibly with *cis*-diols to form cyclic boronate esters (Scheme 6). It is observed that the formation of the ester is more favourable from the anionic boronic acid than from the neutral molecule, i.e., $K_{tet} > K_{trig}$, this difference is typically of up to approximately five orders of magnitude.[65] The acidity constant also varies between the boronate ester and the boronic acid, although it is the ester that proves more acidic, i.e., $pK_a > pK_a'$.[65]

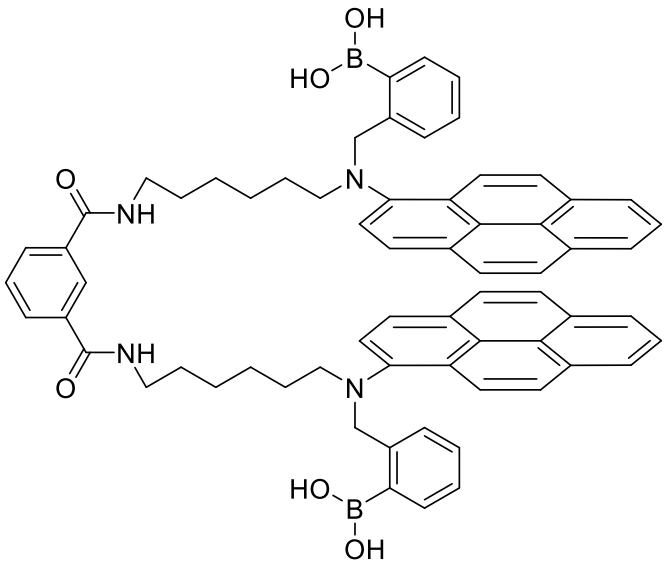
Fluorescence is a useful tool in the design of sensors. As a detection method, fluorescence has a high sensitivity, it being possible to detect a signal from just one molecule. Another benefit is that fluorescence can be measured remotely and the response time is typically under a millisecond.[65]

Figure 12 shows a selection of fluorescent sensors, in these molecules photoinduced electron transfer (PET) quenches the fluorescence when bound to a saccharide or in the free sensor. As a result, the fluorescence of the sensor changes dependent on the presence or concentration of a saccharide.

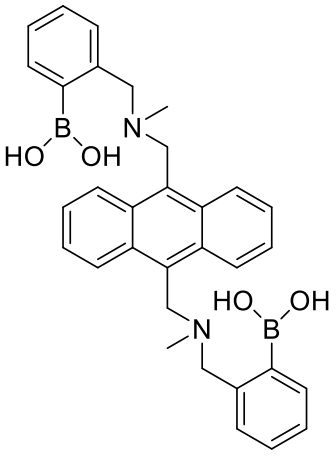
Compound **18** is a pair of molecular tweezers that can bind with a saccharide by either one or both of the boronic acid moieties and forms highly stable complexes with D-glucose and D-fructose (Figure 13).[66] For this molecule, binding with a saccharide quenches the PET and thus an increase of fluorescence is seen upon the binding of the substrate.

Another diboronic acid (**19**), can also bind with saccharides to form a 1:1 complex.[67] As with **18**, fluorescence is enhanced following the binding event, which is again rationalised by the quenching of the PET pathway. In both **18** and **19** this blockage of the PET pathway occurs as a

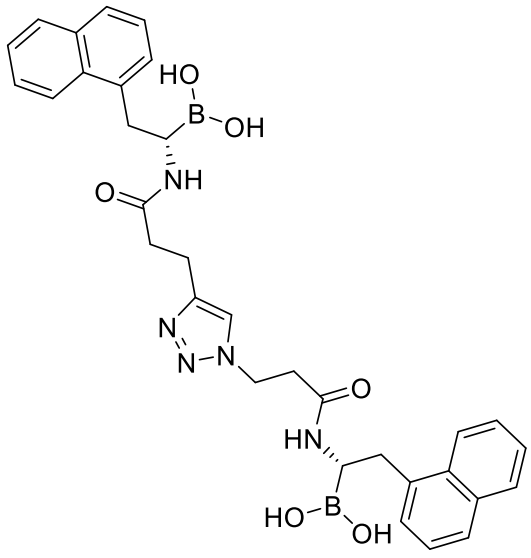
result of a nitrogen-boron interaction that is only possible when the saccharide is bound (Scheme 7).[67] Fluorescence can also be observed by lowering the pH for both **18** and **19**,[66, 67] because of this, the pH range in which the sensor can function effectively is limited. This is a result of the protonation of the nitrogen atom, which prevents PET (Scheme 7).[67] Diboronic acids are inherently selective for monosaccharides but it has been shown by sensor **19** that this selectivity can be tailored towards D-glucose.



18



19



20

Figure 12: Fluorescent saccharide sensors. Molecular tweezers able to bind with either two separate sugars or one sugar molecule with both boronic acid moieties (18),[66] a diboronic acid, which becomes fluorescent when both boronic acids are bound to a sugar, this can be either the same sugar or two different saccharide units (19),[67] and a di-α-amidoboronic acid which strongly binds sugars and exhibits water solubility (20).[68]

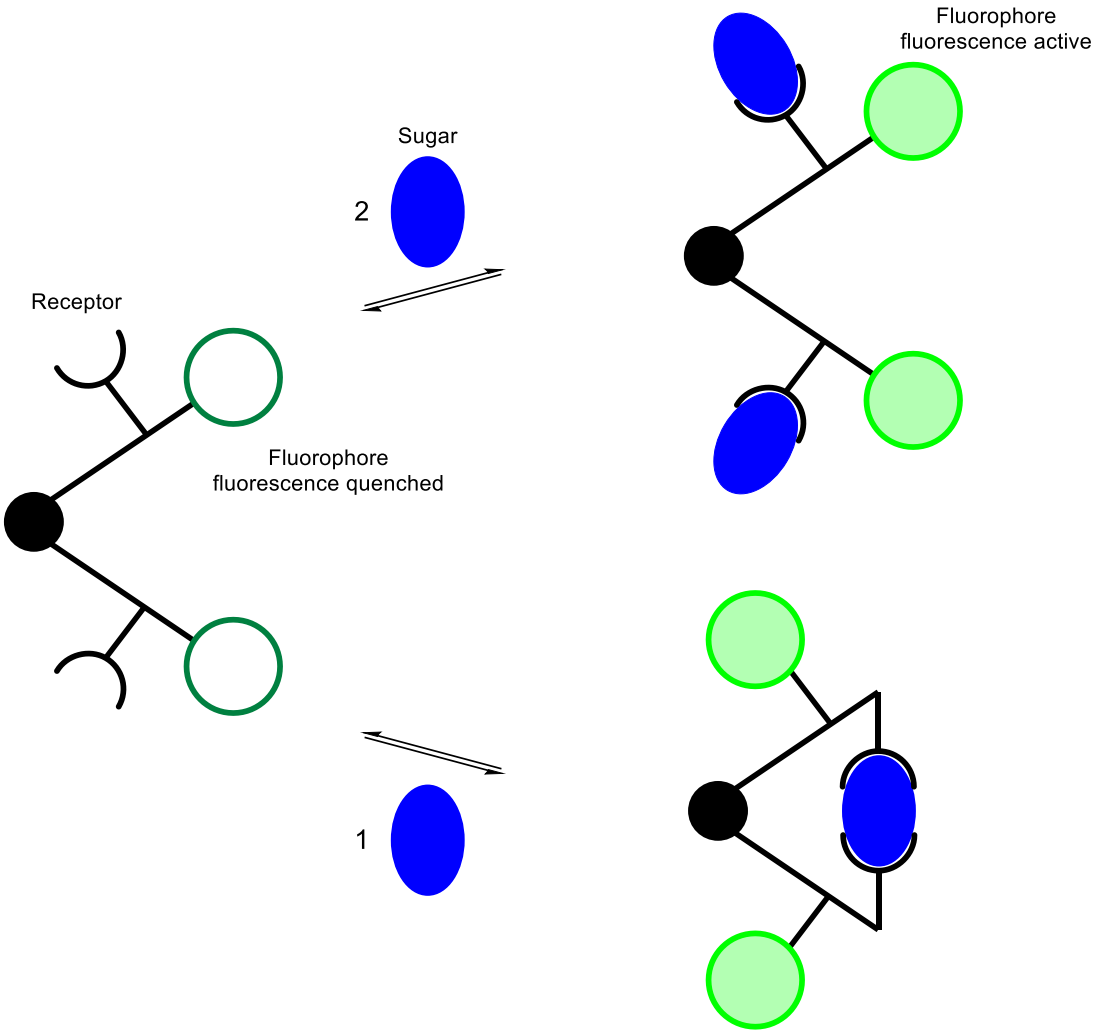
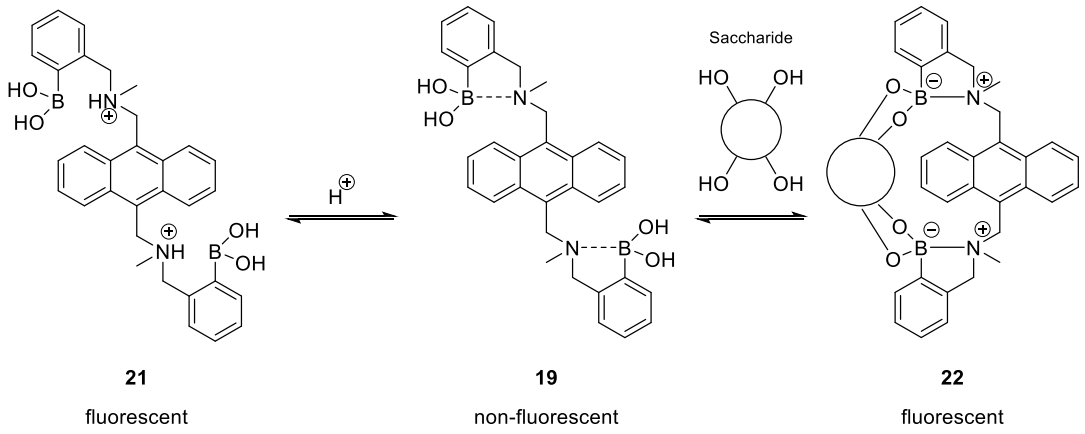


Figure 13: A schematic showing the two possible methods of saccharide binding by the molecular tweezers (18). Binding with either one or two sugar molecules quenches the PET pathway and activates the fluorescence.[66]

The di-α-amidoboronic acid (**20**) is a water soluble sensor for saccharides, for which a decrease in fluorescence was observed when bound to a sugar molecule.[68] The binding constant was compared with the monoboronic acid and was shown to be significantly lower than for the diboronic acid. A computational model was then used to verify the possibility that this is due to the formation of a 1:1 complex between the diboronic acid and a sugar. The model showed that this interaction was possible but further investigation was not carried out.[68]

Besides fluorescence, other detection methods have been used. These include induced circular dichroism (CD), differing absorption of left and right circularly polarized light, wherein the sensor exhibits no CD activity but when bound to a saccharide – which is a chiral molecule – the sensor becomes CD active. Another method is use of sensors that change colour on binding to a saccharide, these are of interest as a strong colour change on detection of a sugar could be used to develop a test paper, similar to pH papers, to indicate the presence of

saccharides.[65] Other methods can involve direct analysis of the sensor following complexation through use of NMR spectroscopy.[69]



Scheme 7: The modes by which fluorescence of 19 can be activated, either by lowering pH to form 21 where the protonation of the nitrogen prevents PET, or by binding with a saccharide to form 22, which allows a nitrogen-boron interaction that blocks PET (formation of a 2:1 saccharide:sensor complex also allows this, fluorescence is activated provided both boronic acids are complexed with a diol).[67]

1.2.3 Detection of AGEs

A number of methods exist to detect and analyse AGEs, most involve invasive sampling followed by the use of analytic techniques.[43] These techniques include high-performance liquid chromatography (HPLC), gas chromatography-mass spectrometry (GC-MS), immunohistochemistry (IHC) staining, boronate affinity chromatography, and gel electrophoresis. A non-invasive method has been developed, which uses an Autofluoresence Reader (AFR) to measure the autofluoresence of skin.[70]

Both HPLC and GC-MS can be used to quantitatively measure the levels of some of the most common AGEs: pentosidine and CML (Scheme 3).[43] Despite the quantitative and specific nature of these methods they have both time and cost restraints, as many steps are necessary to pre-treat samples (reduction and acid hydrolysis of proteins) and the process is expensive. It is also necessary to use a synthesised version of the target AGE as a reference.

IHC can be used to identify regions where AGEs have accumulated in tissue samples.[43] IHC is a method that utilises the targeting of antigens by antibodies to identify the presence, or deficiency, of specific molecules and is frequently used to analyse abnormal cells. There are commercially available antibodies for the identification of a number of AGEs, in particular, pentosidine, CML, CEL, and pyrraline (Scheme 3). CML and CEL are frequently used as targets as they usually have higher concentrations in tissue samples.[43]

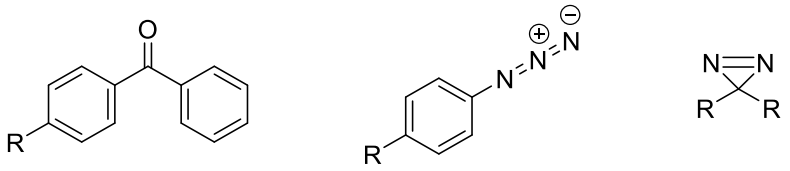
Boronate affinity chromatography draws on the ability of boronic acids to bind with *cis*-diols (Scheme 6). A boronic acid is used as the affinity-ligand in stationary phase, which is able to selectively retain the target molecules.[71] Boronate affinity columns have been used to analyse levels of glycated haemoglobin to assess glycation in diabetic patients along with other glycoproteins.

Gel electrophoresis is used to separate proteins or other large biomolecules, such as strands of deoxyribonucleic acid (DNA), by size. Wells are made in a gel, typically made from agarose, into which the mixture to be separated is placed. A potential difference is then applied across the gel, which causes the charged molecules to migrate through the gel, the molecules are then visualised. A polyacrylamide gel, functionalised with phenylboronic acid moieties, has combined the concepts of boronate affinity chromatography with standard gel electrophoresis and has been investigated for its ability to analyse protein glycation and separate saccharides.[72, 73] These investigations revealed that this gel is capable of separating glycoproteins from diverse mixtures, aiding in their identification.[73]

Measuring the accumulation of AGEs in skin tissue with the AFR exploits the natural fluorescent properties of many AGEs, such as pentosidine.[70] Testing revealed a correlation between skin autofluorescence and the level of accumulation AGEs as measured by analysis of skin biopsies. However, there are limits to this method. Not all AGEs exhibit fluorescence, CML, a common AGE, for example is non-fluorescent, and other fluorophores could be present in the skin tissue.

1.3 Photoaffinity Labelling

1.3.1 Overview of photoaffinity labelling

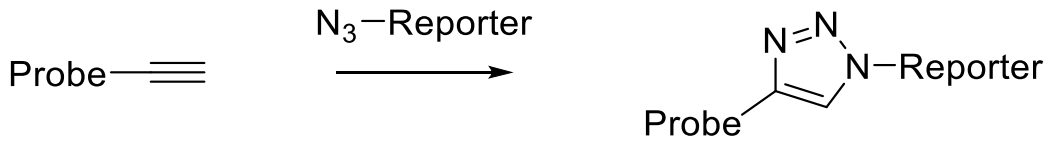


23 **24** **25**

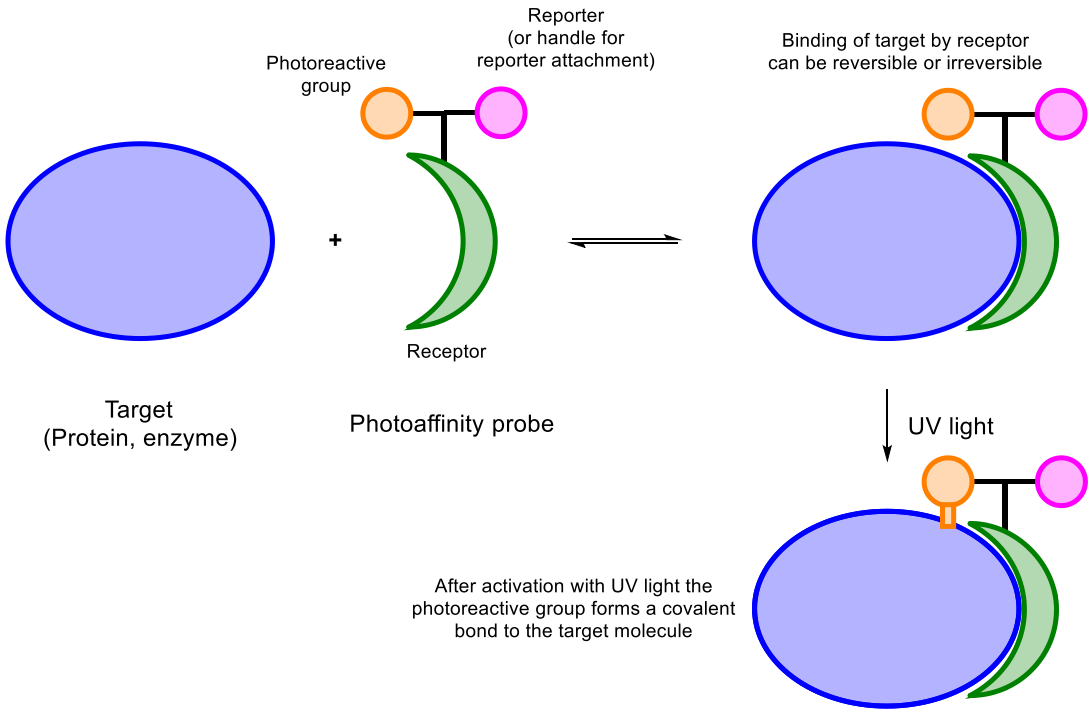
Figure 14: The commonly used photoreactive groups in photoaffinity labelling, a benzophenone (23), an aromatic azide (24), and a diazirine; each of these groups has different benefits and drawbacks pertaining to the ease of synthesis, interruption of receptor-substrate binding, or biological tolerance of the activation wavelength.[74, 75]

Photoaffinity labelling is a technique used to investigate the structure and function of biomolecules. It has also been used to investigate the properties of enzymes, in particular to identify targets, selectivity, and elucidate the three-dimensional structure of the binding site in

its active state, i.e., while bound to a substrate.[74] A photoaffinity probe consists of three basic components, a receptor – or recognition unit – that can bind with the target compound, or family of targets, either reversibly or irreversibly; a photoreactive group, which will form a covalent bond with the target on irradiation with ultraviolet light; and a reporter that can allow for identification or separation of the tagged compound. The mechanism of action of a photoaffinity probe is detailed below in Scheme 9. In some cases the reporter group interferes with binding of the target, this problem has been solved with the use of so-called "click" reactions. This involves the attachment of a functionalisation handle to the probe, following the irradiation step this handle is then used to attach the tag. However, this handle has to be inert toward biological systems, commonly used handles are terminal alkynes, which can react with azides (Scheme 8); this process is referred to as tandem photoaffinity labelling.[75]



Scheme 8: A Huisgen Cycloaddition, an example of a "click" reaction, between an azide and a terminal alkyne, this reaction can be used once the photoaffinity probe has been irradiated and a colvalent link between the target and the probe has been formed. This allows addition of a potential bulky tag once the recognition event has already occurred, reducing potential disruption of the receptor-substrate interaction.



Scheme 9: This schematic diagram shows the mode of activity of a photoaffinity probe. After initial binding to the target, which may be reversible or irreversible, irradiation with ultraviolet light forms a reactive intermediate from the photoreactive group, which then irreversibly forms a covalent bond with the target. Following this the reporter can be used to identify the target.[74, 75]

Commonly used photoreactive groups are shown in Figure 14, each has advantages and disadvantages. Benzophenones (**23**) are the bulkiest of the photoreactive groups in frequent usage; this can disrupt the binding of the target by the recognition unit and lead to non-specific labelling. However, benzophenones are the most simple to synthesise, with a range of precursors commercially available, and their excitation wavelength is unlikely to be damaging to the biological system under investigation (350-360 nm). Aromatic azides (**24**) however, are smaller than other groups used by virtue of their linear shape, allowing for selective labelling, they are also relatively simple to synthesise from amines. Despite these advantages, the absorption wavelength of azides is significantly smaller (<300 nm) than that of other photoreactive groups and this could be considerably damaging to a biological target, they also suffer stability issues in a number of chemical environments such as strongly basic or acidic, and oxidising or reducing. Finally, diazirines (**25**) absorb ultraviolet light of a similar range to benzophenones, thus avoiding biological damage. As with azides they are non-bulky but are stable in the conditions that azides do not tolerate. However, the complicated synthesis of these functional groups is a drawback for the use of diazirines. [74, 75]

A number of different reporter groups have been used in photoaffinity probes, as with the photoreactive groups, each has its own set of advantages and limitations.

Radioisotopes have been used to great effect, particularly ^{125}I and ^{3}H. ^{125}I is easily incorporated into a probe through iodination of a phenol group, while ^{3}H can be introduced in place of a standard hydrogen atom without impacting on the structure. ^{125}I displays high specificity due to a relatively short half-life, however, tritium possesses a much longer half-life and so longer exposure times are required, resulting in lower specificity. Other advantages of radioisotopes are easy detection, high sensitivity, small size leading to less interference with receptor-substrate binding, and can be used both *in vivo* and *in vitro*. However, radioisotopes can damage biological systems and ^{125}I cannot be stored for long periods of time due to decay.

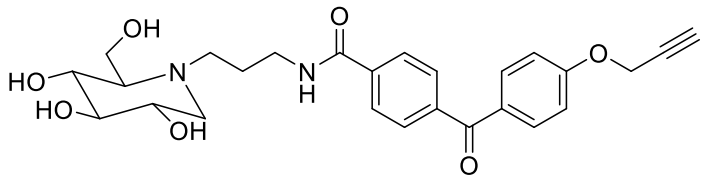
Fluorophores have also be used as reporters in photoaffinity probes. They also provide high sensitivity but are usually bulky and can disrupt targeting. Despite this, the hydrophobic nature of some fluorophores allows for penetration of the cell membrane leading to use *in vivo*.

A third type of reporter, is an affinity tag, which, once attached to a target, can be employed in separation, through affinity chromatography, and detection, through techniques such as IHC staining. Common examples are biotin and epitopes. Similar to fluorophores, these groups are

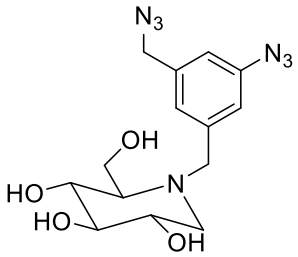
bulky, although this issue can be overcome with the use of tandem photoaffinity labelling, where the reporter is attached to the probe after activation of the photoreactive group.[74, 76]

1.3.2 Use in biological systems

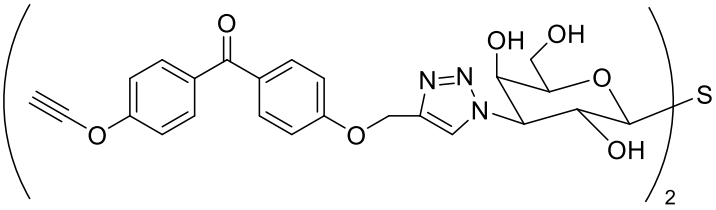
Many studies have utilised photoaffinity probes to investigate a range of proteins and enzymes. Carbohydrate binding proteins have received some interest and a number of probes have been developed to investigate their properties. Structurally, these probes all share a saccharide or saccharide-like moiety, which is used as the recognition unit. However, a range of different reporters and photoreactive groups have been used in conjunction with both of these types of receptor.[74, 77]



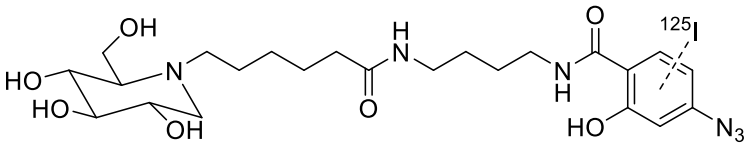
26



27



28



29

Figure 15: Photoaffinity probes used to investigate carbohydrate binding proteins. Both 26 and 28 use a benzophenone photoreactive group, while 27 and 29 use an aryl azide. Only 29 has the reporter group, the radioisotope ^{125}I, attached to the probe, the others utilise tandem photoaffinity labelling to attach the reporter after binding of the target molecule.[74, 75, 78-81]

Figure 15 shows some recently reported examples of photoaffinity probes for carbohydrate binding proteins. **26** is a label for glucosidases, enzymes that catalyse the hydrolytic cleavage of glycosidic bonds, specifically β-glucosidases, which target terminal glucose units of oligosaccharides bound by a β-glycosidic bond.[78] The recognition unit of this label is the

iminosugar, 1-deoxynojirimycin, which has shown previous use in photoaffinity labels; as this iminosugar has a stronger affinity with the target, it was used to overcome the interference of the bulky benzophenone group on binding to the enzyme. This probe was shown to have high specificity for the target enzyme, being able to selectively label it from a complex mixture of proteins, however, the efficiency of the label was low. 1-deoxynojirimycin was also used as the recognition unit for **27**; it was shown that this label could specifically target α-glucosidases and exhibited no interaction with other glycoside hydrolase enzymes, including β-glucosidases.[80] The aryl azide gave the label high specificity and efficiency and it was shown that **27** could selectively label α-glucosidases from a diverse proteome. The benzyl azide survives the irradiation step and it was shown that this could be successfully coupled with a reporter group following the covalent binding of the label to the target enzyme. Another tandem photoaffinity probe, **28**, uses thiodigalactoside to selectively target galactin-3, a protein that binds to β-galactoside. The heterocyclic linkage in label **28** provides rigidity, which is thought to be instrumental in the high selectivity observed in complex mixtures, but leads to diminished label efficiency, despite this it was shown that **28** could label galactin-3 even at low concentrations. Unlike the other probes shown, **29** uses the radioisotope ^{125}I as a reporter, although like **27** selectively targets α-glucosidases through use of 1-deoxynojirimycin as a receptor. As with other probes discussed **29** is highly specific for its target and is able to selectively label it within a mixture of proteins. These probes are powerful tools for the investigation of protein and enzyme activity, structure and function.

1.4 Summary

Glycoproteins are common biomolecules which serve a wide range important functions. They are characterised by the presence of glycans, sugar moieties, attached during or shortly after their synthesis. These glycans can be essential to the function and structure of the protein and have been linked to various properties and functions.

Non-enzymatic glycosylation, or glycation, or proteins, however, is detrimental to their function and can cause denaturation. Glycation occurs when a protein reacts with the open chain form of a reducing sugar. While in an equilibrium, these sugars predominantly exist as either a five or six membered ring and so glycation is more commonly observed in those with high blood sugar levels, such as patients with diabetes mellitus. Glycated proteins can undergo further transformations to produce advanced glycation end products.

AGEs have been linked to the onset and progress of a number of diseases, which have been detailed here. There are a number of different ways in which AGEs promote the progress of

these diseases: the formation of cross-linked proteins can lead to the formation of amyloids, oxidation of AGEs produces reactive oxygen species that can cause further protein damage, and the interaction of AGEs with the receptor for advanced glycation end products can cause chronic inflammation. With such a strong impact on health, analysis and detection of both AGEs and glycated proteins is an important target for research. Current methods for analysis include high pressure liquid chromatography, gas chromatography mass spectrometry, immunohistochemistry staining, electrophoresis, skin autofluoresence, and boronate affinity chromatography. While these existing methods can be effective there are drawbacks such as time, cost, and limited range of the AGEs they can detect.

Molecular recognition relies upon a highly selective interaction between a receptor and a substrate. Examples of this phenomenon exist in nature but also in manufactured molecular sensors. Molecular sensors for the detection of carbohydrates have been reported in the literature, many of which utilise the equilibrium between boronic acids and boronate esters to target sugar molecules.

A technique that employs molecular recognition is photoaffinity labelling. This process requires a probe possessing: a recognition unit for targeting; a photoactive group, which upon activation will bind irreversibly to the target; and a reporter to identify compounds that have been labelled.

This thesis will combine existing technologies for the recognition of sugar molecules with photoaffinity labelling, with the aim of developing a probe for the detection of glycated proteins.

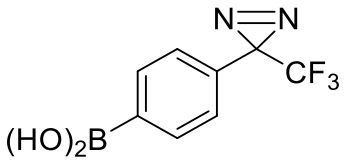
Chapter 2: Photoaffinity Probe Synthesis

It was proposed that a boronic acid could be utilised as a recognition unit in a photoaffinity probe for the analysis of glycated proteins. This was suggested as the use of boronic acids as sugar sensors has been widely reported in the literature.[82-84] As discussed, the detection of glycated proteins is desirable as they, and their reaction products, have been identified as important biomarkers in the onset and progress of various diseases.[15, 22, 24, 27-29] The design and synthesis of such a molecule is discussed within this chapter.

2.1 Designing the Photoaffinity Probe

The first task was to design a photoaffinity probe that possessed the required properties. As discussed, a photoaffinity probe consists of three main components: a receptor, a photoaffinity group, and a reporter.

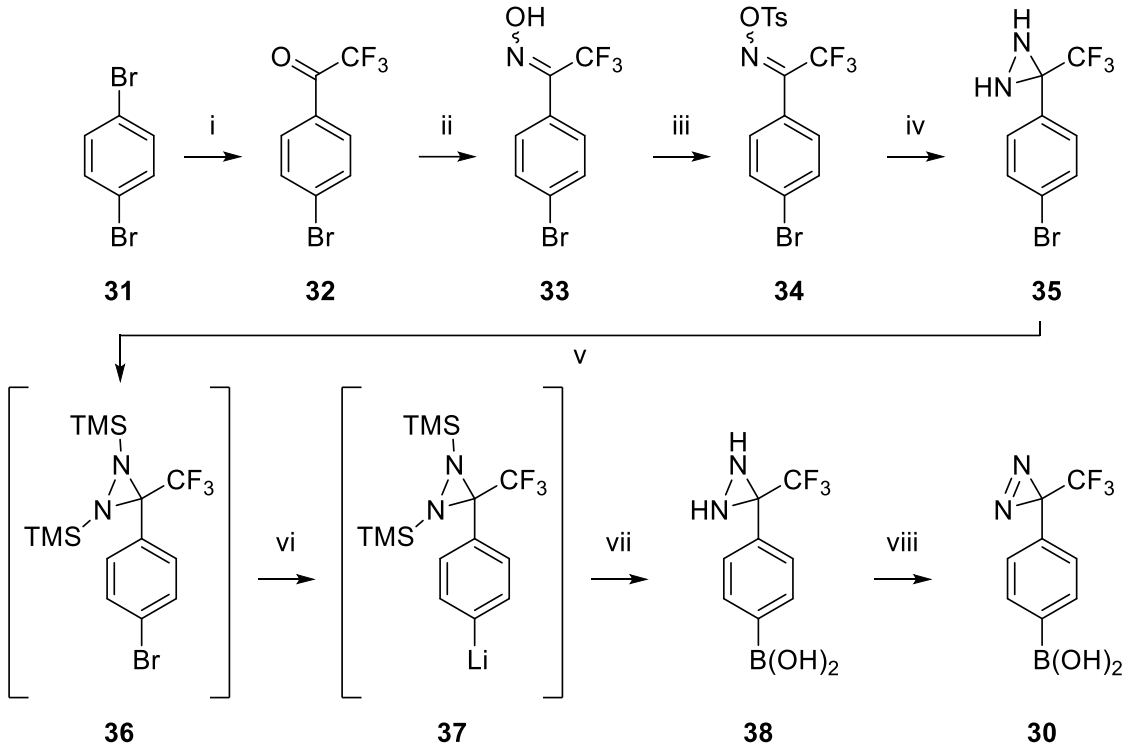
The receptor for this probe would be a boronic acid, as specified by aim of this study. A diazirine was selected for use as the photoaffinity group. This would provide stability in a wide range of chemical environments; allow the use of a longer wavelength of ultraviolet light to activate thereby reducing the risk of damage to the target protein; and minimise steric interference with the target allowing the probe to bind more easily.[74, 75] It was decided that a specific reporter would not be added to the initial probe to simplify the synthesis and that mass spectrometry would be used to identify labelled proteins. These components were then combined to give the suggested design of the probe as shown in Figure 16 (**30**).



30

Figure 16: (4-(3-(trifluoromethyl)-3*H*-diazirin-3-yl)phenyl) boronic acid (30) the design of the photoaffinity probe.

A literature search led to a reported synthetic route of a similar compound with a carboxylic acid functional group in place of the boronic acid,[85] this pathway could be easily adapted to yield the target compound instead. The reported route has a total of eight synthetic steps and it was identified that this could potentially limit the final yield through the accumulation of even small losses on each step. However, the reported yields[85] of each step are high, mitigating this concern and it was ultimately decided that this would be the most efficient synthesis when also taking into account the cost of starting materials.



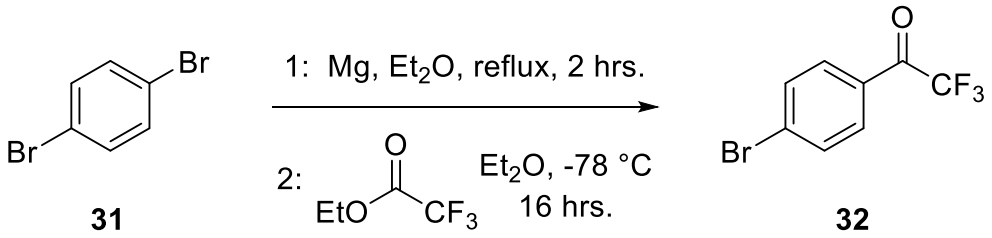
Scheme 10: Devised synthetic pathway to the photo affinity probe, the reagents and conditions for each step are: (i) 1) Mg, Et_2O, reflux, 2 hrs., 2) $EtOC(O)CF_3$, Et_2O, -78°C, 16 hrs.; (ii) $HONH_2{\cdot}HCl$, EtOH:Pyr (1:2), 60 °C, 2 hrs.; (iii) *p*-TsCl, NEt_3, DMAP, DCM, 0 °C → r.t., 3 hrs.; (iv) $NH_{3(l)}$, DCM, -78 °C, 18 hrs.; (v) TMSOTf, NEt_3, DCM, -78 °C → r.t., 4 hrs.; (vi) n-BuLi, hexane, -78 °C, 1 hr.; (vii) $B(OMe)_3$, hexane, -78 °C → r.t.; (viii) I_2, NEt_3, MeOH, r.t., 30 mins.[85, 86]

The resulting synthetic route used to produce **30** is shown in Scheme 10. The conditions and reagents for steps **i - vi**, and **viii** are as previously reported, with step **viii** following the described method but with a different starting material.[85] Step **vii** follows a previously reported method used to synthesise boronic acids.[86]

2.2 Synthesis of the Photoaffinity Probe

Once the design of the probe and the synthetic route had been finalised, the project could progress onto the next stage and work on the synthesis itself started. In this section the individual steps of the synthesis are detailed and the results of the reactions are discussed.

2.2.1 Synthesis of 4'-bromo-2,2,2-trifluoroacetophenone



Scheme 11: Preparation of 4'-bromo-2,2,2-trifluoroacetophenone.

This reaction (Scheme 11) was done a number of times; initially, the reaction was performed on a small scale as a trial run before repeating on a larger scale to build up a supply of the 4'-bromo-2,2,2-trifluoroacetophenone for later reactions. Table 2 shows the range of yields, and the scale of different runs of the reaction; in general, the larger scale reactions afforded a better yield, rationalised as the amount product lost by either competing reactions (i.e. Grignard reagent with residual moisture) or, incomplete extraction from the crude by distillation, makes up a smaller proportion of the total.

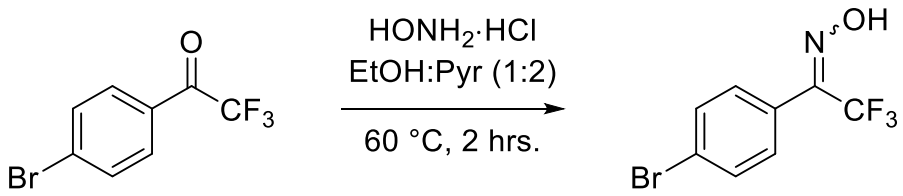
These experiments followed the published method.[85] However, following initially poor yields (experiments 1 and 2 from Table 2) the method for the generation of the Grignard reagent was optimised by grinding the magnesium turnings under nitrogen prior to the reaction to produce activated magnesium and adding a catalytic amount of iodine, which improved yields.

Table 2: The yields obtained from different attempts of the reaction shown in Scheme 11.

Experiment	Scale (mmol of 31)	Yield (mmol) (percentage yield in brackets)
1	21.20	4.00 (19 %)
2	63.50	25.77 (41 %)
3	106.0	91.77 (87 %)
4	106.0	63.68 (60 %)
5	212.0	130.1 (61 %)
6	212.0	151.0 (71 %)

When necessary, the product was purified by vacuum distillation, the boiling point was observed to be 60°C at 0.4 mbar. ^{1}H NMR and ^{13}C NMR spectroscopy of the distillate revealed the pure product. In some experiments it was not possible to fully separate the unreacted 1,4-dibromobenzene (**31)** from the product (**32**); however, it was deemed that this would not interfere with the following step of the synthesis so the sample was used without further purification.

2.2.2 Synthesis of 1-(4-bromophenyl)-2,2,2-trifluoroethanone oxime



32 **33**

Scheme 12: Preparation of 1-(4-bromophenyl)-2,2,2-trifluoroethanone oxime.

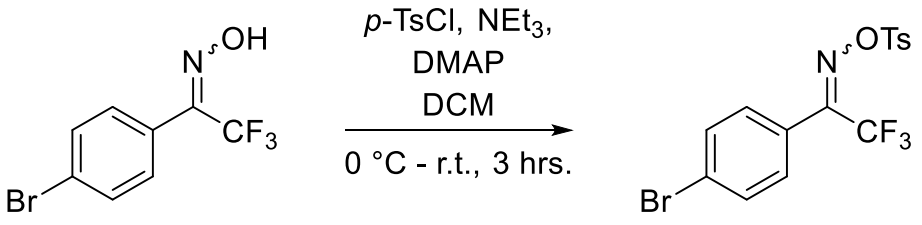
Synthesis of the oxime **33** was performed following the reported method (Scheme 12).[85] This reaction generally proceeded without issues and gave reasonable yields comparable to those reported.[85] However, some experiments exhibited lower than expected yields, in particular Experiment 2 from Table 3. A suggested explanation for this is the presence of unreacted 1,4-dibromobenzene in the starting material, causing an overestimation of the predicted yield. The large difference observed in Experiment 2, however, is greater than would be expected for this alone, suggesting other factors, such as experimental error, may have been involved.

Table 3: The yields obtained from different attempts of the reaction shown in Scheme 12.

Experiment	Scale (mmol of 32)	Calculated yield (mmol) (percentage yield in brackets)
1	1.98	1.64 (83 %)
2	25.7	3.47 (14 %)
3	91.7	56.0 (61 %)
4	63.7	60.2 (95 %)
5	118.5	72.1 (61 %)
6	146.2	68.51 (47 %)

The product was purified by column chromatography for the smaller scale reactions; this was not suitable for larger quantities of crude, so instead the product was deprotonated with sodium hydroxide and dissolved in aqueous solution. This was then washed with diethyl ether before neutralisation of the aqueous layer with a 1.0 M solution of hydrochloric acid. The product was then extracted from the aqueous phase with diethyl ether, which was subsequently washed with water and brine; this method gave lower yields (see Experiments 3, 5 and 6 in Table 3) but was easier to perform on the larger scale.

2.2.3 Synthesis of 1-(4-bromophenyl)-2,2,2-trifluoroethanone O-(*p*-tolylsulfonyl) oxime



33 **34**

Scheme 13: Preparation of 1-(4-bromophenyl)-2,2,2-trifluoroethanone o-(*p*-tolylsulfonyl) oxime.

As for previous reactions, preparation of **34** followed the method reported.[85] However, as show in Table 4, the first attempts at the reaction gave impure products and very poor yields

following recrystallisation and so the *p*-toluenesulfonyl chloride used was analysed and it was found to have partially degraded into *p*-toluenesulfonic acid, which would act to revert **33** into **32** by hydrolysis.

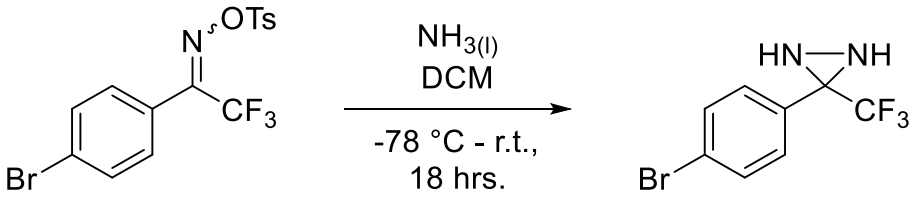
The *p*-toluenesulfonyl chloride was then purified by recrystallisation and the reaction repeated, this gave improved, but still poor, yields (experiments 3 and 4 from Table 4). As such a new source of *p*-toluenesulfonyl chloride was purchased before the reaction was repeated. Following attempts at the reaction in Scheme 13 showed acceptable yields and the product was sufficiently pure for further reactions.

As the recrystallisation of **34** yielded large crystals, which appeared to be of good quality, it was suggested that single crystal X-ray diffraction could be used to elucidate the stereochemistry of the oxime. Crystals were submitted for in house X-ray analysis (performed by Dr Gabriele Kociok-Köhn), but despite a number of attempts they did not produce a diffraction pattern and a structure could not be obtained.

Table 4: The yields obtained from different attempts of the reaction in Scheme 13.

Experiment	Scale (mmol of 33)	Calculated yield (mmol) (percentage yield in brackets)
1	1.87	0.90 (48 %)
2	11.6	0.67 (6 %)
3	55.96	12.3 (22 %)
4	66.16	36.4 (55 %)
5	63.43	45.8 (73 %)
6	67.16	52.01 (77 %)

2.2.4 Synthesis of 3-(4-bromophenyl)-3-(trifluoromethyl) diaziridine

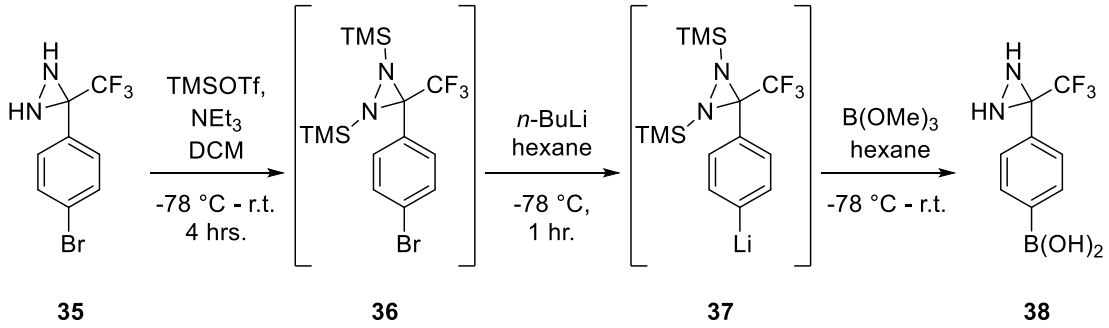


34 **35**

Scheme 14: Preparation of 3-(4-bromophenyl)-3-(trifluoromethyl) diaziridine.

This reaction was performed using the materials described by Bender *et al*, [85] however, an autoclave was not used as it was suggested it would be simpler to introduce the liquid ammonia with a dry ice condenser. Each time this reaction was performed, the target compound was isolated in high yields (84 % - 98 %) and purity.

2.2.5 Synthesis of (4-(3-(trifluoromethyl)diaziridin-3-yl)phenyl)boronic acid

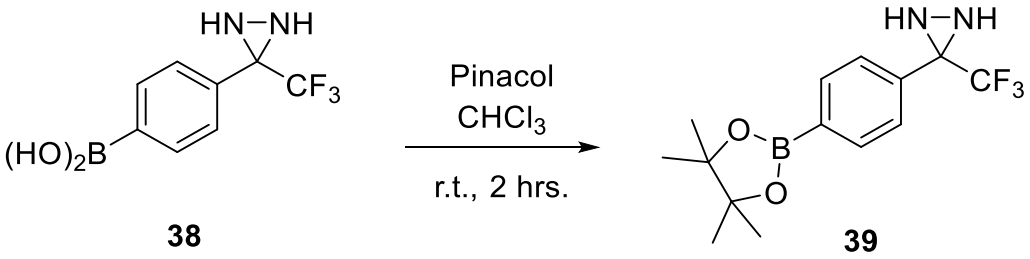


Scheme 15: Preparation of (4-(3-(trifluoromethyl)diaziridin-3-yl)phenyl)boronic acid.

The syntheses of the protected diaziridine **36** and the organolithium compound **37**, were performed according to the reported method,[85] all reactions were performed using standard Schlenk techniques. Due to their instability towards moisture neither **36** nor **37** were isolated and were instead immediately used in the next step to produce **38**. During the purification of the boronic acid, **38**, some problems were encountered. During column chromatography, it was found that the compound formed streaks on a silica gel column, suggested to be a result of the acidic protons. To alleviate this issue it was decided to form the pinacol ester, **39**, to protect the acidic protons and simplify the purification of the compound.

An initial small scale trial of this reaction gave a low yield of **38** (38 %); however, subsequent larger scale reactions gave improved yields (67 % and 98 %), although, as these measurements were taken on the crude product after only a simple work-up, the impurities present will have adjusted these figures to be higher than the true yield of the desired compound.

2.2.6 Synthesis of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)-3H-diaziridine



Scheme 16: Preparation of 41.

The pinacol ester was prepared by adapting a previously reported method (Scheme 16).[87] As intended, this allowed for better purification. Yields from this reaction varied as shown in

Table 5, this is most likely to reflect upon the purity of the starting material rather than the efficiency of the reaction.

Table 5: The yields obtained from different attempts of the reaction in Scheme 16.

Experiment	Scale (mmol of 38)	Calculated yield (mmol) (percentage yield in brackets)
1	2.82	1.66 (59 %)
2	12.4	12.1 (97 %)

The pinacol ester, **39**, was not reverted to the boronic acid, **38**, after purification as it was deemed to be unnecessary for the function of the probe. As there is a equilibrium between the ester and the boronic acid, it was proposed that, in the presence of a glycated protein, **39** would undergo transesterification and form an ester with sugar groups on the target protein. It has historically been shown that boronic esters formed with sugars are more energetically favourable than those formed with other diols;[88] this supports the hypothesis that the pinacol ester, **39**, would be just as able to target a glycated protein as the boronic acid, **38**.

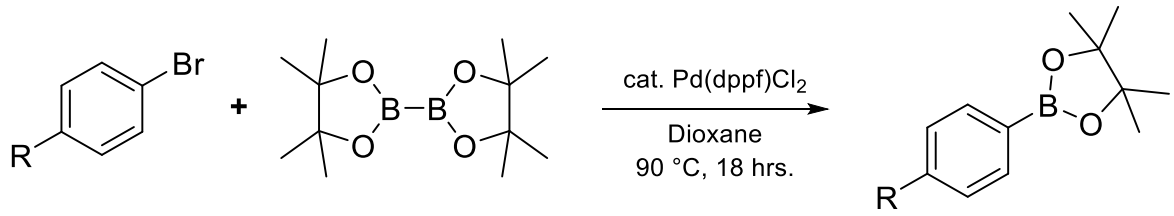
It was decided that the probe should be stored in the form of the pinacol ester **39** and that the diazirine would only be generated immediately prior to use due to its photosensitivity.

2.3 Optimisation of Photoaffinity Probe Synthesis

Following the successful synthesis of **39**, it was of interest to refine the synthetic pathway so that it would be simpler and more efficient to create more of the photoaffinity probe in the future. The protection of the diaziridine and subsequent addition of the boronic acid were highlighted as the parts of the synthesis with the greatest potential for improvement. These steps were also some of the more challenging synthetically due to the moisture sensitivity of the compounds. As a result, it was decided to focus the effort to improve the synthesis on these steps.

In work by Ishiyama, Murata, and Miyaura,[89] it has been shown that the boronate pinacol ester can be added directly to a target molecule without first forming the boronic acid by a Miyaura reaction (Scheme 17). Altering the synthesis to employ this method would remove the need to form the protected diaziridine (**36**) and the organolithium complex (**37**) used in the original synthesis. The addition of the boronate ester could theoretically be performed at any point in the synthetic route so it was necessary to investigate which of these reactions would be the most efficient.

2.3.1 Investigation of Boronate Ester Synthesis using a Miyaura reaction



Scheme 17: Borylation of a para-substituted aryl halide by a Miyaura reaction.

Each of the following reactions was initially performed on a small scale and analysed with NMR to assess their efficiency. These reactions were performed using the conditions given in Scheme 17, which was based on a previously reported method.[90]

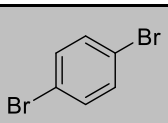
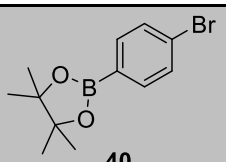
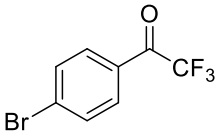
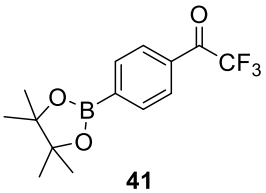
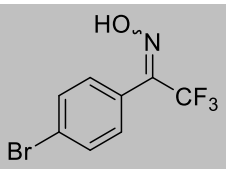
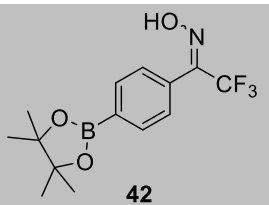
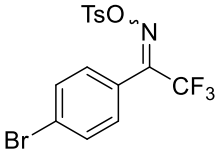
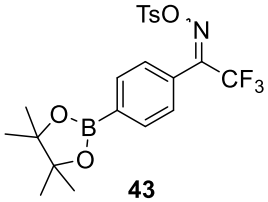
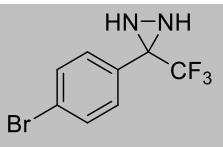
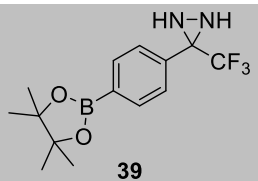
Determination of the conversion rates was performed by NMR analysis of the crude product. To simplfy the NMR, the reaction mixture was first filtered through silica with *tert*-butyl methyl ether (tBME) and then washing with water in order to remove the catalyst and some of the reactants. It was decided to add dichloromethane (DCM) as an internal standard to the prepared NMR sample, this was because no signals from the products shown in Table 6 would appear in the range in which the DCM peak appears in ^{1}H NMR. The standard was added in a ratio of 6:1 to the predicted yield of the product, such that the standard peak and that of the pinacol methyl groups would have the same integration at 100 % conversion.

Table 6 shows the results of the investigation into the efficiency of the Miyaura reaction at different stages along the synthetic route to the target compound (**39**). The highest conversion rates were observed for experiments 2, 4 and 5, the reactions of compounds **32**, **34**, and **35**. Therefore, these would be the optimal points to incorporate this borylation method into the synthesis of **39**.

The low conversion of 1,4-dibromobenzene (**31**), Experiment 1 in Table 6, can be explained by the potential for double substitution introducing the possibility of extra impurities in the product. The same catalyst and reactants in similar conditions to those used in these experiments have been previously used to replace both halides on a dihalogenated arene with aryl groups by a Suzuki coupling,[91] thus giving weight to the possibility that diborylation could be a cause of a lower yield in this situation. Further, a literature search revealed that, while examples exist of single borylation of dihalogenated aromatic compounds, these tend to go via an organolithium compound.[92-94] Reactions such as these inherently limit the potential for double borylation as the formation of a doubly lithiated compound is highly unfavourable.

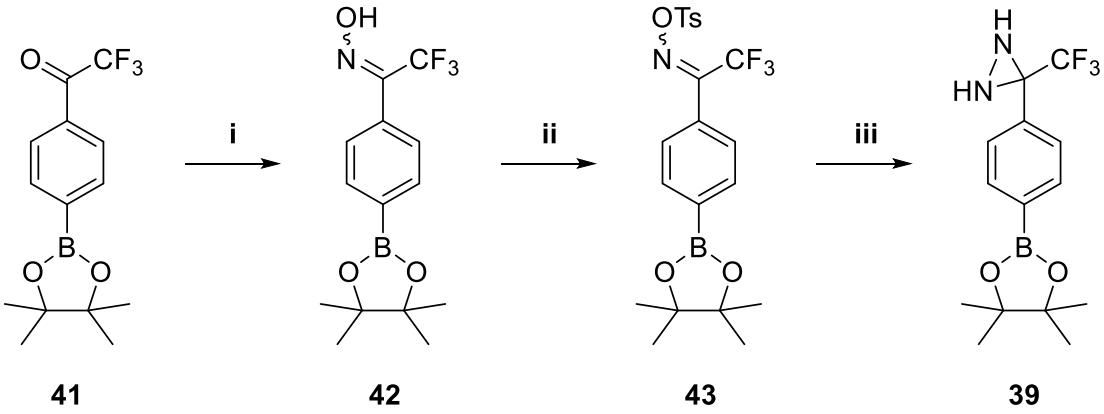
Experiment 3 from Table 6, the conversion of the oxime **33** to its analogous boronate ester **42**, also shows a low yield. This was thought to be a result of an interaction arising between the oxime and the palladium catalyst. Other studies have used an oxime-palladium complex as a catalyst for other Suzuki-Miyaura coupling reactions.[95, 96] This behaviour could introduce the potential for a similar complex to form with oxime **33** or even the product of the reaction, **42**. In the formation of such a complex, the oxime would replace the labile chloride ligands and thereby poison the catalyst, ultimately stunting the progress of the reaction. This activity would in turn result in lower than expected yields.

Table 6: The percentage conversion of aryl halides, from the synthetic route used to generate the probe (Scheme 10), to the pinacol ester analogues using a Miyaura borylation (Scheme 17). * Conversion rates calculated from NMR data (Appendix I) using DCM as an internal standard.

Experiment	Starting Material	Product	Conversion* (%)
1	 **31**	 **40**	25
2	 **32**	 **41**	84
3	 **33**	 **42**	24
4	 **34**	 **43**	54
5	 **35**	 **39**	86

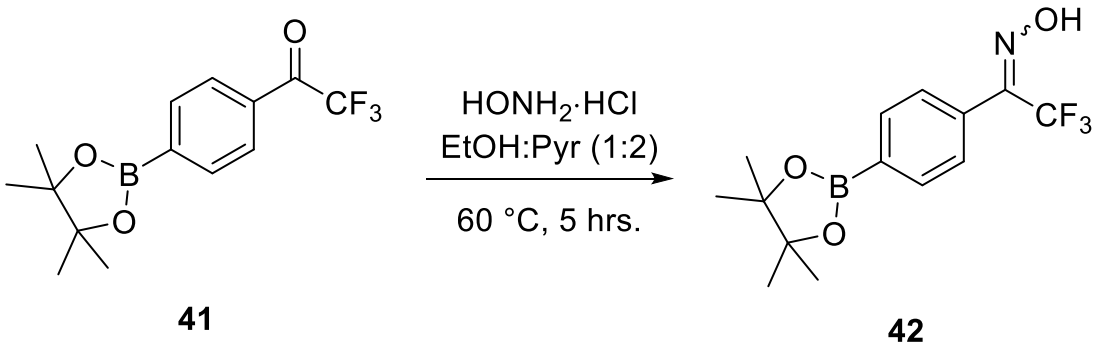
Following these results, the feasibility of incorporating the Miyaura borylation into the synthetic route was investigated by synthesising **39** from the borylated ketone, **41**, following

the pathway shown in Scheme 18. It was decided to use **41** as the starting material as this would cover the synthetic route to **39** from each of the three highest yielding Miyaura reactions. These investigations would be necessary as any increase in yield obtained from the improved borylation method could be offset by losses in other reactions along the synthetic pathway.



Scheme 18: The proposed synthetic route to 41 **from the products of borylation by a Miyaura reaction. Reaction conditions and reagents are as used in Scheme 10: (i) $HONH_2$·HCl, EtOH:Pyr (1:2), 60 °C, 2 hrs.; (ii) *p*-TsCl, NEt_3, DMAP, DCM, 0 °C → r.t., 3 hrs.; (iii) $NH_{3(l)}$, DCM, -78 °C, 18 hrs.[85]**

2.3.2 Synthesis of 2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one oxime

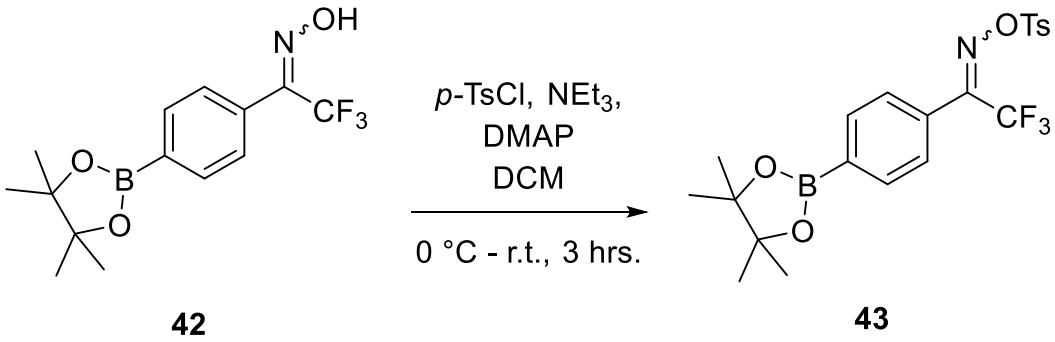


Scheme 19: Preparartion of 2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one oxime.

It was decided to follow the procedure used in the synthesis of oxime, **33**, from **32** (Scheme 12) in the synthesis of **42**. After two hours had elapsed, the time required for the formation of **33**, the reaction mixture was analysed by thin layer chromatography. The plate was visualised with a curcumin dip in order to stain compounds containing a boronic ester. Two spots were observed, one of which with the same R_f value as the starting material, so the reaction was continued for another three hours to allow it to progress to completion.

Further analysis after the additional three hours of reaction time showed that all starting material had been consumed. The product was then isolated and purified by column chromatography to yield the target compound, **42**, with an 87 % yield. This is similar to the yields obtained from the conversion of **32** (83 %) (Table 3), when performed on a similar scale to that used in this reaction (0.5 mmol of **42**).

2.3.3 Synthesis of 2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one *O*-tosyl oxime

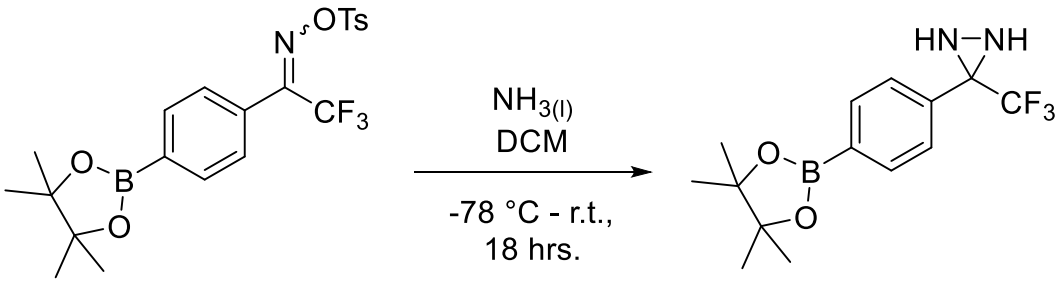


42 **43**

Scheme 20: Preparation of 2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one *O*-tosyl oxime.

Synthesis of **43** (Scheme 20) proceeded by the same method used in the formation of **36** (Scheme 13). Following the reaction, the product was purified by recrystallisation from chloroform to give the target compound in high purity with a yield of 82 %, comparable to the yields obtained in the formation of **34**.

2.3.4 Synthesis of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine



43 **39**

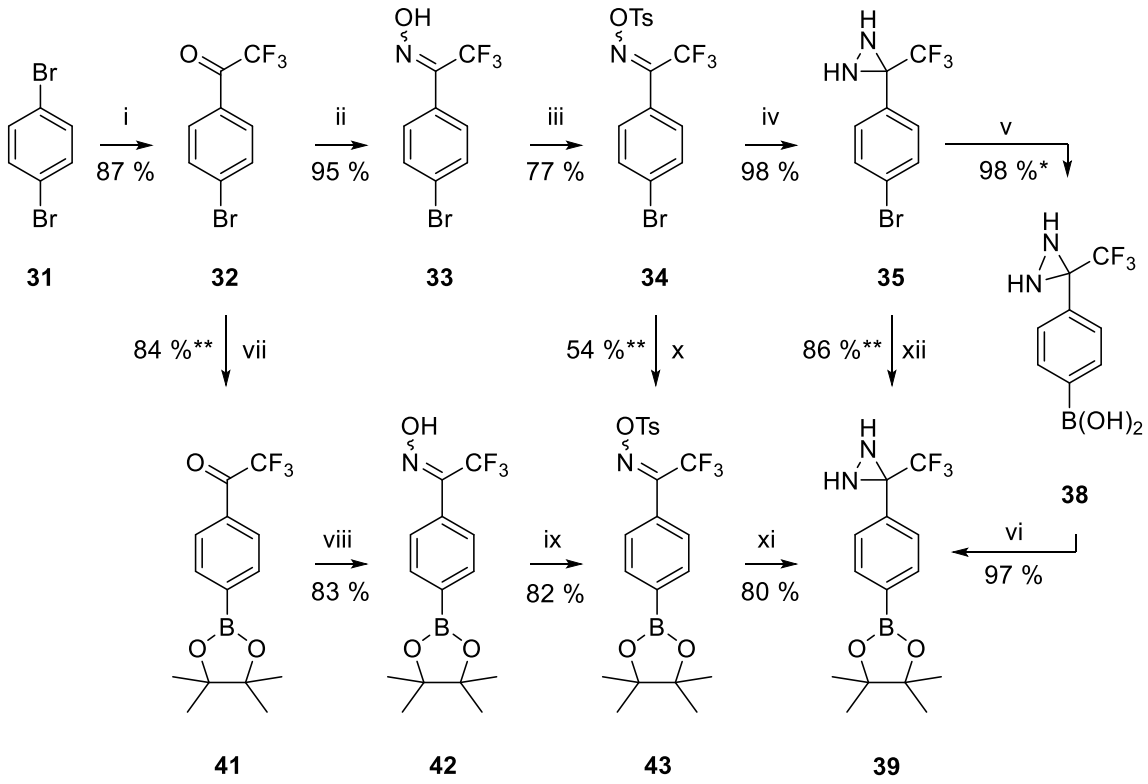
Scheme 21: Preparation of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl) diaziridine.

The formation of the diaziridine **39** was performed using the same reported method as used in the previous synthesis of **35**.[97] After the completion of the reaction, the product was purified

by column chromatography to give **39** with a yield of 80 %. This is a good yield and is close to the lower end of the range of yields observed for the formation of diaziridine **35**, it also compares well to the yield obtained when forming **39** using the Suzuki coupling (Table 6) suggesting that either could be used in an alternative synthetic route to the photoaffinity probe.

2.4 Conclusions

The target compound, **39**, has been successfully synthesised and a number of possible synthetic routes investigated. The results of the initial synthesis and the subsequent investigation into possible optimisations are shown below in Scheme 22. Only the three Suzuki-Miyuara reactions with the highest conversion rates are shown (steps vii, x, and xii), as the rates of the other two reaction were too low to be considered viable for the synthetic pathway.



Scheme 22: Yields obtains from different routes to the target compound of 39**, unless otherwise specified, the values show represent the highest yield obtained for that reaction amongst all experiments performed. * Value obtained from crude yield. ** Values obtained from conversion rates. Reagents and conditions are: (i) 1) Mg, Et2O, reflux, 2 hrs., 2) EtOC(O)CF3, Et2O, -78°C, 16 hrs.; (ii) HONH2•HCl, EtOH:Pyr (1:2), 60 °C, 2 hrs.; (iii) *p*-TsCl, NEt3, DMAP, DCM, 0 °C → r.t., 3 hrs.; (iv) NH3(l), DCM, -78 °C, 18 hrs.; (v) 1) TMSOTf, NEt3, DCM, -78 °C → r.t., 4 hrs., 2) n-BuLi, hexane, -78 °C, 1 hr.; (vii) B(OMe)¬3, hexane, -78 °C → r.t.; (vii) cat. Pd(dppf)Cl_2, Dioxane, 90 °C, 18 hrs.; (viii) HONH2•HCl, EtOH:Pyr (1:2), 60 °C, 2 hrs.; (ix) *p*-TsCl, NEt3, DMAP, DCM, 0 °C → r.t., 3 hrs.; (x) cat. Pd(dppf)Cl_2, Dioxane, 90 °C, 18 hrs.; (xi) NH3(l), DCM, -78 °C, 18 hrs.82; (xii) cat. Pd(dppf)Cl_2, Dioxane, 90 °C, 18 hrs.**[85, 90]

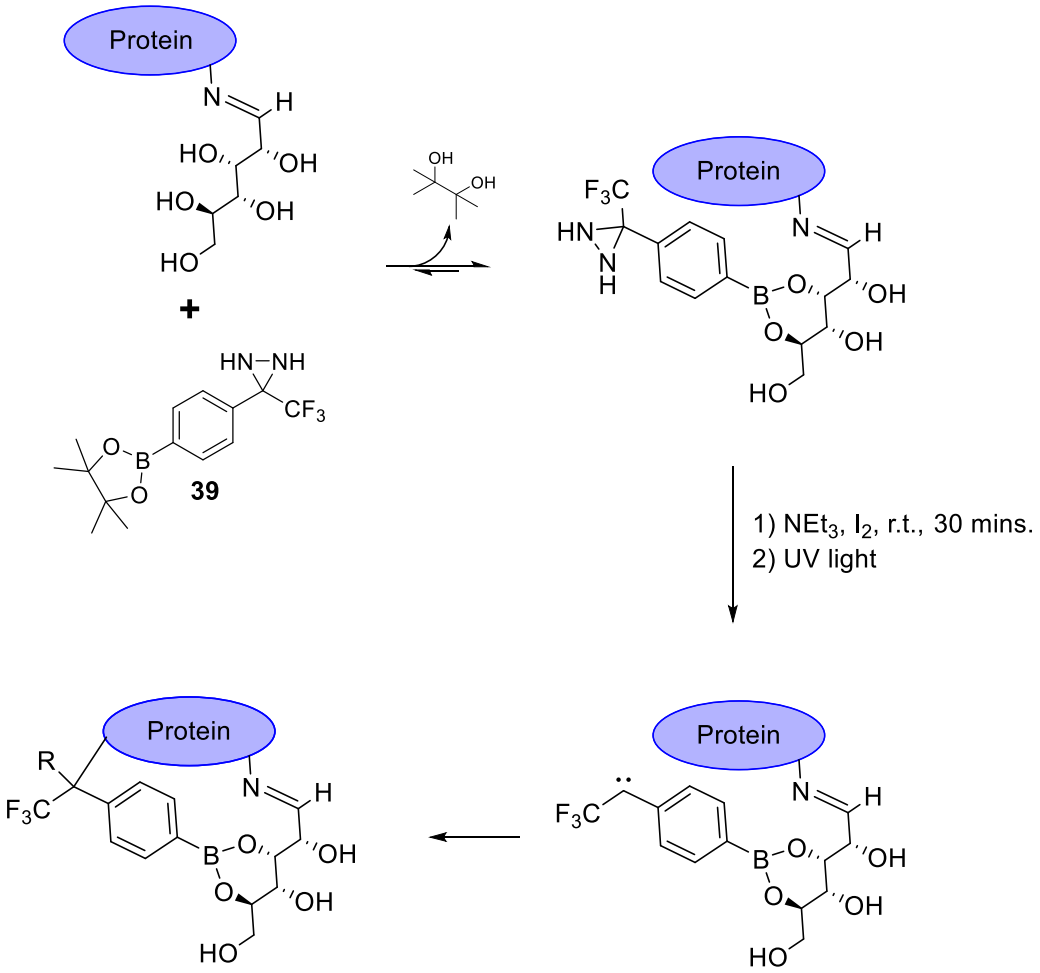
Considering the yield of each step, the originally planned synthesis would appear to be the most efficient with an overall yield of 59 %. However, this uses the yield of step v which is taken from the crude yield as further purification was not performed before the following reaction. Given the likely presence of impurities this value would be an overestimate of the true yield, furthermore, a wide range of crude yields were measured for this reaction (38 % up to 98 %) suggesting limited repeatability, possibly due to the sensitivity of the intermediate compounds.

Taking this into consideration, the next highest overall yield (53 %) is obtained by the route which follows the original synthesis but diverges in the final step by using the Suzuki-Miyuara reaction to form the boronate ester (steps i, ii, iii, iv, and xii). While the value shown for the Suzuki-Miyuara reaction is based on a conversion obtained from analysis by nuclear magnetic resonance spectroscopy, it is not unreasonable to expect that the isolable yield would be close to this value and so is a good estimate of the possible yield. In light of these results, this synthetic route is a viable alternative to the originally proposed pathway.

The alternative route splitting at the second step (steps i, vii, viii, ix, and xi) has an overall yield of 39 %. However, this route takes the steps viii, ix, and xi, each of which were performed fewer times than the corresponding reactions on the halogenated analogues (steps ii, iii, and iv) and it is possible that further experiments could show higher yields for these reactions. Additionally, the conditions used in steps vii, ix, and xi, are the same as those used in steps ii, iii, and iv respectively. These were conditions used in the original synthesis going via compounds **32**, **33**, **34**, and **35**,[85] and, as such, are likely to have been optimised for these compounds. Therefore, further investigation could be done to optimise the conditions for steps viii, ix, and xi for their starting materials and improve the overall yield of this alternative pathway.

Chapter 3: Protein Labelling

With the synthesis of the probe completed, the focus of the project moved to investigate the potential of this molecule, **39**, for use in labelling glycated proteins. These investigations were performed on a range of proteins, the results of which are discussed within this chapter.



Scheme 23: Labelling of a generic glycated protein by 39**, following cross-esterification of** 39 **and the glycated protein, an oxidation is performed[85] to produce the diazirine, which is then irradiated with ultraviolet (UV) light to form a carbene. The carbene then inserts across a bond in the protein.**

The proposed mechanism for the labelling of a protein by **39** is shown in Scheme 23. As discussed previously, boronic acids will favourably form esters with sugar moieties,[65, 88] this is applied here in the equilibrium between the probe and a glycated protein. It is proposed that a more favourable boronic ester will form with the sugar residue on the protein than with the pinacol molecule, this acts as the targeting mechanism of the photoaffinity probe.

Once the association between **39** and the protein has been formed, the diaziridine is oxidised; the method shown in Scheme 23 is that reported in the synthesis followed to create **39**.[85] This would then be followed by irradiation with ultraviolet (UV) light to convert the diazirine into a

carbene. As the protein and probe are now closely positioned, the highly reactive carbene species is likely to insert across a bond in the protein forming an irreversible covalent bond between the probe and the target protein. Following this covalent linkage, the labelled protein can be analysed without risk of the probe and protein becoming detached.

3.1 Human Serum Albumin

It was decided to perform these tests initially using human serum albumin (HSA) (Figure 17) as a target for the probe.

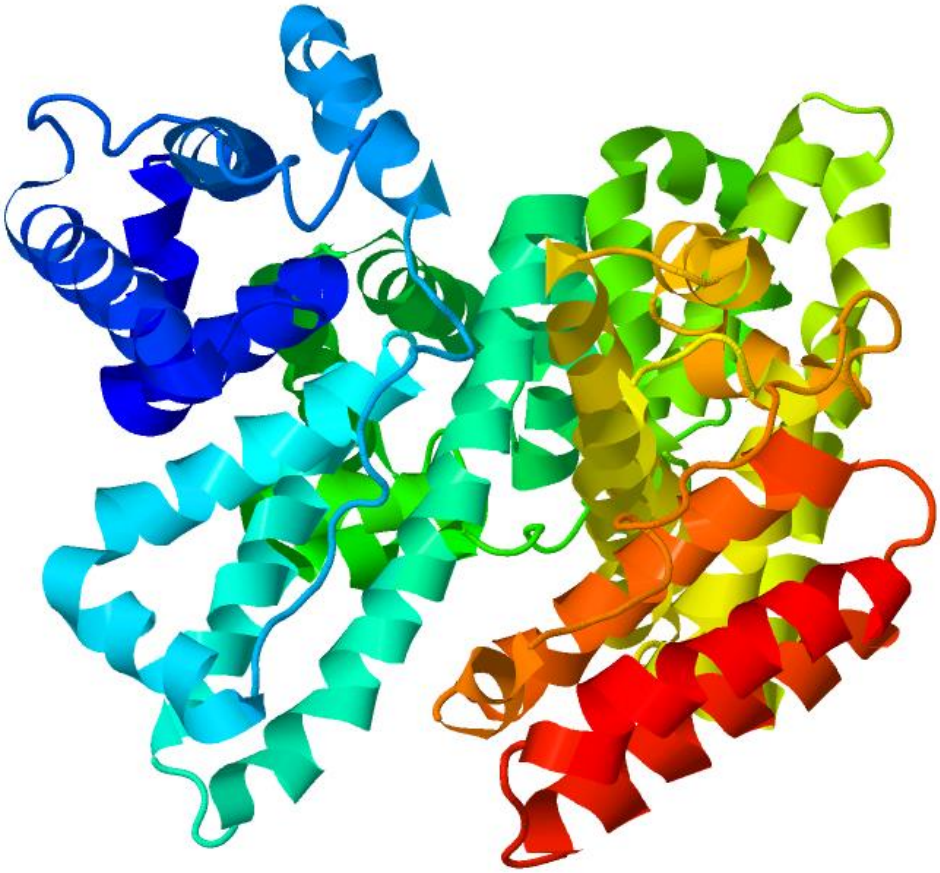


Figure 17: The structure of a free molecule of human serum albumin[98] (PDB ID: 4K2C, image generated by Jmol version 14.6.1). HSA is comprised of three domains (I, II, and III) which are linked by helical extensions.[99]

HSA is the most prevalent of plasma proteins, it constitutes around 60 % of the total amount of protein found in serum and plays a role in a number of biological processes such as: regulating blood pH, mediation of lipid metabolism, and the transport of a number of biological molecules and drugs.[99, 100] It has been shown that HSA is commonly glycated, especially in diabetic patients, and measurements of glycated HSA have been used for many years to measure the control of blood sugar levels.[99] It was due to these factors and its availability that HSA was chosen as the first protein on which to test the photoaffinity probe.

It was decided to first investigate forming the diazirine *in situ* by performing the oxidation once **39** had been mixed with the protein. It was believed this would be advantageous as the boronic acid would have theoretically already formed a complex with the protein before the generation of the photosensitive compound. This would reduce the possibility of the diazirine converting to a carbene prior to the formation of an association complex between the probe and the target protein, which may occur if the diazirine was formed before combining the compounds.

To achieve this, a solution of HSA was combined with a solution of **39**. It was decided to perform the experiment with two different volumes of the solution of **39** to give molar ratios of 5.0 and 10.0 of **39** to HSA. The aim of this was to investigate the optimum amount of **39** required to reliably and detectably label the protein. These solutions of HSA and **39** were stirred together overnight to facilitate the formation of a boronate ester between **39** and the sugar moieties of the glycated HSA. Iodine and triethylamine were then added to oxidise the probe and produce the diazirine (Scheme 23). This oxidation was performed in the dark to avoid prematurely causing the diazirine to decompose. Once sufficient time had passed for the reaction to go to completion, the solution was exposed to ultraviolet light with a wavelength of 354 nm to reveal the carbene. These solutions were then analysed by mass spectrometry to determine if the HSA samples had been successfully labelled.

Table 7: Results of labelling HSA with 39**. *Expected increase is calculated by multiplying the molar ratio of** 39 **added to the sample by the molecular weight of the fragment of** 39 **expected to be linked with the protein following the labelling process. This was taken to be** 39 **with the loss of pinacol and nitrogen (165.91 gmol^{-1}) as shown attached to the glycated protein in the final step of Scheme 23.**

Sample	Molar ratio of 39	Recorded Mass (gmol^{-1})	Difference from unlabelled (gmol^{-1})	Expected Difference (gmol^{-1})*
Unlabelled HSA	0	66469.86	0.00	0.00
Labelled I	5	67317.03	+ 847.17	+ 829.55
Labelled II	10	68495.19	+ 2025.22	+ 1659.21

As shown in Table 6, the increase in molar ratio gave a direct increase in mass detected. The increase from the unlabelled sample is approximately equal to the expected increase calculated from of the number of equivalents of the probe added for five equivalents of the probe. However, a larger than expected increase is observed for ten equivalents. This is possibly a result of the method used to calculate the expected increase. The mass per unit of probe added was assumed to be 165.91 gmol^{-1}, this was derived from the probe as a boronic ester as shown in Scheme 23, if it is based on boronic acid form, a figure of 201.94 gmol^{-1} per

equivalent of probe is reached instead. This would put the expected increase in the same region as that observed. It is possible that a greater loading of probe would lead to a change in the 3D configuration of the protein following labelling, causing the boronic ester to no longer be favourable. Another explanation for this discrepancy comes from the observations that for a higher equivalence of the probe, the signal in the mass spectrum became weaker. This observation is likely to be due to a larger range of possible products from labelling as some HSA molecules could remain unlabelled while others would then accumulate a larger amount of the probe, leading in turn to a wide range of masses.

From these results, it was decided that 5.0 equivalents of **39** would be used for further experiments as it gave a detectable increase in mass whilst maintaining a good signal strength.

These results have shown that **39** can successfully label glycated HSA. Following this success, it was decided to further investigate the mechanism by which the probe acted to determine if the molecule behaved as believed (Scheme 23).

In order to show this, it was decided to use two control experiments where molecules with structures similar to **39** but with certain functionality removed were used in place of the probe. The first of these was a "non-targeting" control, wherein a molecule that possessed the photoactive group but lacked the boronate ester was used. This was to show that the formation of a boronate ester was essential to the labelling. To mimic the structure of **39** as closely as possible, it was decided to use **35** for this experiment with the same procedure described above. Secondly, a "non-coupling" control, using a variation of the probe with a boronate ester moiety but without the carbene in order to show that the formation of an ester alone was not sufficient to keep the probe and protein coupled throughout the analysis. It was decided to test this using **39** but by altering the procedure to no longer include the oxidation step.

Table 8: Mass spectrometry results from control experiments used to investigate the labelling mechanism. In both tests the compound used was added in a 5:1 molar ratio of compound:protein, with HSA used as the target protein. *Labelling procedure performed without the oxidation step.

Sample	Compound Added as Probe	Recorded Mass ($gmol^{-1}$)
Unlabelled HSA	-	66469.86
Control I	**35**	66559.24
Control II	**39***	66561.46

In both control tests, no significant increase is shown over the unlabelled HSA as was the expected result. This shows that the boronic acid moiety is indeed acting to target the glycated

protein, bringing the two molecules together before the photoactive group can bind them irreversibly, however the formation of the boronic ester alone was not sufficient to maintain the labelling of HSA through the analysis.

3.2 Casein

Following the successful result with HSA, it was decided to investigate other proteins. The first of these selected was casein in part due it its availability. Casein is a major component of several types of milk, accounting for between 20 % and 45 % of the total protein content in human milk at different stages and an average of 80 % of the protein content of bovine milk.[101, 102] In milk, the casein micelle is constructed of four subunits, three phosphoproteins (αS_1-, αS_2-, and β-casein) and one phosphoglycoprotein (κ-casein).[102] The latter has been shown to undergo glycation and react further to form AGEs. [103] As a result, it was decided to use isolated κ-casein to further test the labelling ability of the molecular probe, **39**.

The labelling procedure for HSA was again followed with casein. However, in this case the experiments met with mixed results; analysis by mass spectrometry on the initial attempt gave no evidence of any large molecules, showing several peaks at relatively low masses of less than 8000 $gmol^{-1}$ (Labelled I, Table 9). It was suggested that this could be a result of over exposure to UV light causing the degradation of the protein. It has been shown exposure to UV light can induce changes in the structure and activity of casein proteins, promote oxidation, and lead to a reduction in concentration against non-irradiated samples.[104-107]

To counter this issue, the labelling was reattempted with a reduced exposure time, in this instance a range of peaks were observed in the mass spectrometry analysis, however the results still showed an average reduction in mass (Labelled II, Table 9).

Table 9: Results of mass spectrometry analysis after performing the labelling procedure on κ-casein. 39 **was added in a 5:1 molar ratio of probe:protein, oxidised with an I_2/NEt_3, before exposure to UV light (354 nm) for the indicated time. Expected mass difference was calculated using the same method described in Table 7.**

Sample	UV Exposure Time (s)	Recorded Mass ($gmol^{-1}$)	Difference from unlabelled ($gmol^{-1}$)	Expected Difference ($gmol^{-1}$)
Unlabelled Casein	0	19035.51	0.00	0.00
Labelled I	300	7945.18	- 11090.33	+ 829.55
Labelled II	60	17954.67	- 1080.84	+ 829.55

The loss of mass is much smaller with the shorter exposure time and this would suggest that observed changes in mass are a result of the ultraviolet radiation. From these results it was

decided that this method of labelling and analysis was not suitable for use with κ-casein as it would be difficult to conclusively show that the protein had been labelled with only by mass spectrometry. As the mass spectrum data of the labelled samples shows a wide range of peaks compared to the single sharp peak in the unlabelled data, it was suggested that there was a large amount of variation in the mass lost. This would lead to difficulties predicting the mass loss from UV irradiation even at lower exposure times and would inevitably disguise the true extent to which the target protein had been labelled and analysis of the results would be inaccurate. Additionally, a literature search showed no previous studies in which casein had been labelled with a photoaffinity probe, further suggesting the protein is not suitable for this method of analysis.

3.3 Macrophage Migration Inhibitory Factor

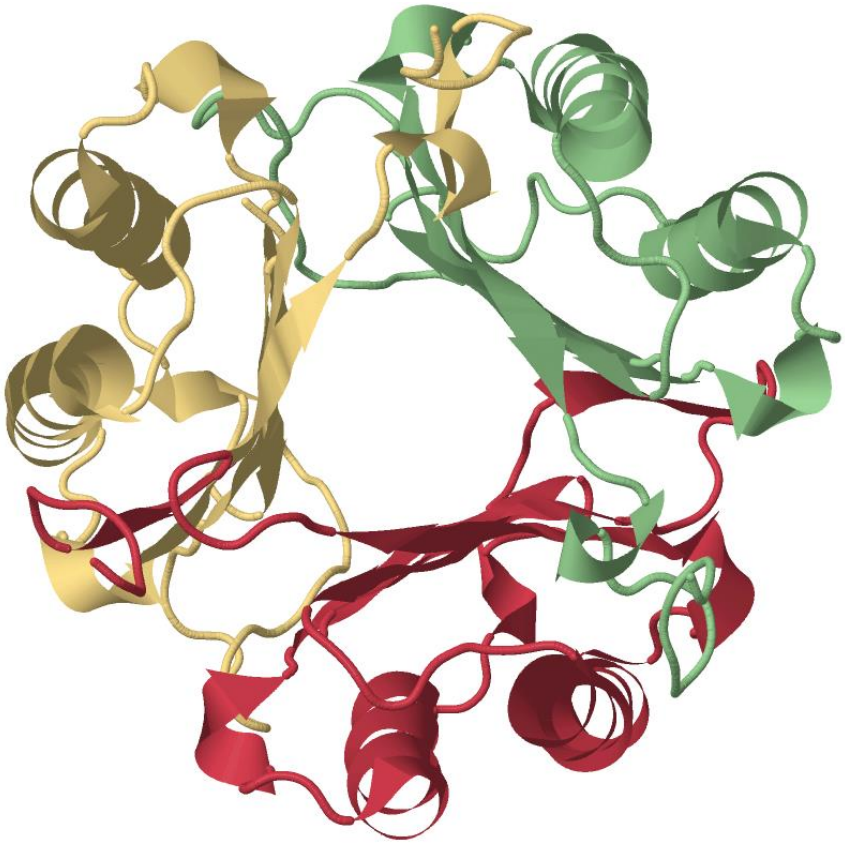


Figure 18: The crystal structure of macrophage migration inhibitory factor[108] (PDB ID: 1MIF, image generated by Jmol version 14.6.1). In this orientation, the three monomers can be seen at the top left, top right, and bottom surrounding the solvent channel at the centre.

It was then decided to investigate another protein, in this case, macrophage migration inhibitory factor (MIF) was selected. As shown in Figure 18, MIF exists as a trimer that forms due to the intertwined chains and interaction of β-sheets at the interface between two monomers.[108]

MIF has been shown to have a range of functions within a biological system. It has an important role within the immune system. MIF is produced by T-cells and is necessary for their activation, while also being pivotal in initiating inflammatory response and being a necessary component for immune responses.[109, 110] Overexpression, and in some cases underexpression, of MIF has been shown to have a role in a wide range of diseases including: cancer,[109, 110] obesity and diabetes related diseases,[111, 112] kidney disease,[113] neurological disease and disorders,[114-116] and various inflammatory diseases.[111, 117]

In addition, modified glycated MIF has been shown to be present in Alzheimer's disease brains,[116] this led to the decision to select MIF as a good target for testing the photoaffinity probe. The labelling was performed using the previously described method.

Table 10: Results of mass spectrometry analysis on samples obtained by performing the labelling procedure previously described on macrophage migration inhibitory factor. 39 **was added in a 5:1 (probe:protein) molar ratio, oxidised with I_2/NEt_3, then irradiated with UV light (354 nm) for 60 seconds. Expected difference was calculated by assuming five units of** 39 **would be attached (as described in Table 7).**

Sample	Recorded Mass ($gmol^{-1}$)	Difference From Unlabelled ($gmol^{-1}$)	Expected Difference ($gmol^{-1}$)
Glycated MIF	12123.25	0.00	0.00
Labelled MIF I	23721.51	+ 11598.26	+ 829.55
Labelled MIF II	12787.15	+ 663.9	+ 829.55

The initial labelling of MIF led to an unexpected result (Labelled MIF I, Table 10) with a much higher observed increase in mass than expected. It was proposed that this increase was a result of the formation of a dimer following the labelling process. The increase is of a similar amount expected for the addition of another molecule of the protein, however it is lower than would be predicted. This is unexpected given that the addition of the label to both protein molecules should put the resultant mass higher than would be expected for an unmodified dimer. This discrepancy was suggested to be caused by further modification of the protein, such as oxidation, which would be possible given that the addition of iodine and triethylamine to oxidise the probe was performed *in situ*. It was unfortunately not possible to perform addition analysis to either confirm or disprove this explanation as the protein rapidly became insoluble in water, although this observation could support the suggestion of further modifications and altered proteins are known to become insoluble.[24]

It was decided to repeat the experiment in order to allow for further analysis and to confirm the result, however, the second labelling led to a different outcome (Labelled MIF II, Table 10). In this case, it was observed that the recorded mass was lower than expected but did show an

increase over the unlabelled glycated MIF. The detailed mass spectrometry of the labelled sample showed a range of peaks (Table 11).

Table 11: Details of all peaks with greatest intensity observed in the mass spectrometry analysis of the product from the second labelling of macrophage migration inhibitory factor (Labelled MIF II, Table 10). * The number of probe molecules that the increase from the unmodified protein is equivalent to was calculated using the assumption of 165.91 $gmol^{-1}$ added per probe molecule as used previously.

Sample	Recorded Mass ($gmol^{-1}$)	Difference From Unlabelled ($gmol^{-1}$)	Equivalent number of probe molecules*
Glycated MIF	12123.25	0.00	0.00
Peak I	12168.14	+ 44.89	0.27
Peak II	12266.61	+ 143.36	0.86
Peak III	12380.69	+ 257.44	1.55
Peak IV	12568.28	+ 445.03	2.68
Peak V	12787.15	+ 663.90	4.00

From these data, the progression in the mass of each of the observed peaks can be seen to be of a loosely similar amount to that of one additional probe molecule. The peaks in Table 11 would therefore roughly equate to the addition of zero (Peak I), one (Peak II), two (Peak III), three (Peak IV), and four (Peak V) probe molecules. The values calculated in the table however, have, in some cases, substantial variation from the addition of an integer number of probe molecules. This could be a result of the different possible conformations of the molecules of **39** added to the protein, as discussed previously. Also, a small portion of the variation may arise a result of slight modifications to the protein structure that can be caused as a result of the ionisation process.[118] The progression of masses observed, alongside the overall mass increases detected, support that the probe molecule has successfully labelled the glycated MIF protein.

As the probe was added in a 5:1 probe:protein ratio, it would expected that the highest peak would arise from the addition of five probe molecules as was seen for HSA. This discrepancy could be a result of modifications to the protein as it was suggested were present in the possible dimer observed in the first labelling (Labelled MIF I, Table 10). If the mass of the protein was reduced by the modifications as this interpretation of data in Table 11 suggests, then this would provide an explanation as to why the mass of the possible dimer is seen to be lower than would be expected.

Following these experiments, it was decided to repeat the labelling procedure on a sample of non-glycated MIF. This was of interest as it would enable further investigation into the selectivity of the photoaffinity probe by allowing for a direct comparison between a glycated

and non-glycated version of the same protein. As shown in Table 12, there is a negligible difference in the mass of non-glycated MIF observed before and after the labelling procedure. This alongside the previous control tests with HSA, confirms that the boronic ester allows **39** to selectively target glycated proteins.

Table 12: Results of mass spectrometry analysis on samples obtained by performing the labelling procedure on non-glycated macrophage migration inhibitory factor. 39 **was added in a 5:1 (probe:protein) molar ratio, oxidised with I_2/NEt_3, then irradiated with UV light (354 nm) for 60 seconds.**

Sample	Recorded Mass ($gmol^{-1}$)	Difference From Unlabelled ($gmol^{-1}$)
Unlabelled MIF	12757.81	0.00
Non-glycated Control	12757.61	-0.20

3.4 Electrophoresis

Whilst mass spectrometry allows for accurate analysis of the labelled samples, it was of interest to investigate alternative methods that would allow for simpler detection of proteins labelled with 41. In this case, it was decided to look at the possiblity of using gel electrophoresis, a commonly used tool for analysing and separating protein samples.

It has been shown that a polyacrylamide gel functionalised with boronic acid moieties can be used to separate glycated proteins with electrophoresis.[72, 73] It was suggested that a similar procedure could be used to instead detect the presence of a boronic acid, in the form of the photoaffinity probe, attached to a protein. In the earlier study, the boronic acid was incorporated into the solid phase and a diol, in the form of a sugar, attached to the protein. Therefore it was suggested that incorporating a diol into the solid phase could allow separation of the labelled proteins, giving a simple way to confirm the presence of the label in a protein sample (Figure 19). The labelled protein band could then be cut out for further analysis, such as possible identification of specific glycation sites.

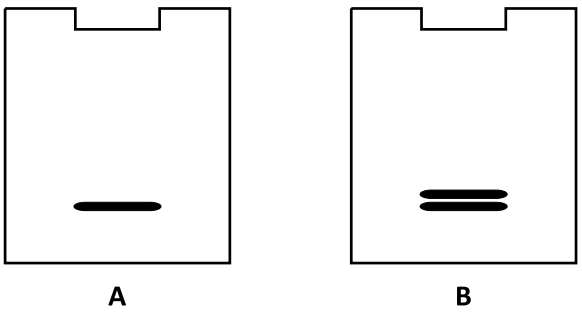
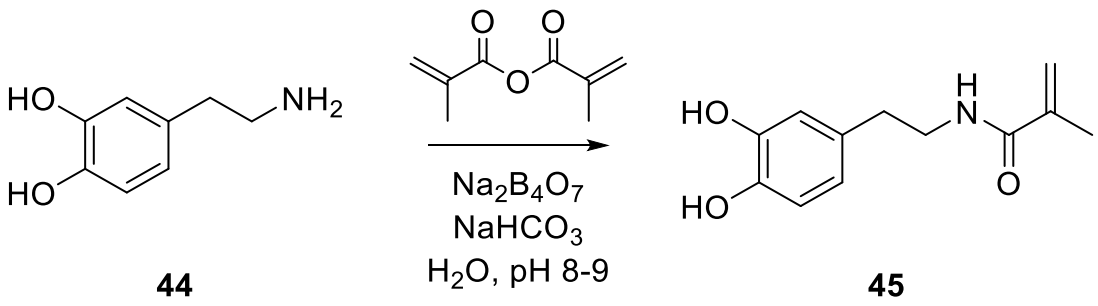


Figure 19: Comparison of a glycated protein on regular polyacrylamide gel (A) with one functionalised with a boronic acid (B), the second band visible in B is a result of the sugar interacting with the boronic acid. Comparing a regular gel with one functionalised with a diol, would theoretically lead to the same differences observed for a sample of a protein labelled with 39**.**

To create a polyacrylamide functionalised with a diol, a compound with both a diol and an acrylamide moiety had to be designed and synthesised. It was decided to use dopamine (**44**) as the starting point in this synthesis as it possesses a diol and the amine could be used to simply add the acrylamide moiety. A literature search yielded a previously reported method for the of synthesis N-[2-(3,4-Dihydroxy-phenyl)-ethyl]acrylamide (**45**) (Scheme 24).[119]



44 **45**

Scheme 24: Preparation of *N*-(3,4-dihydroxyphenethyl)acrylamide.

The procedure was followed without any modifications and after recrystalisation of the crude product as described in the method, the target compound, **45**, was obtained in high purity with 60.3 % yield.

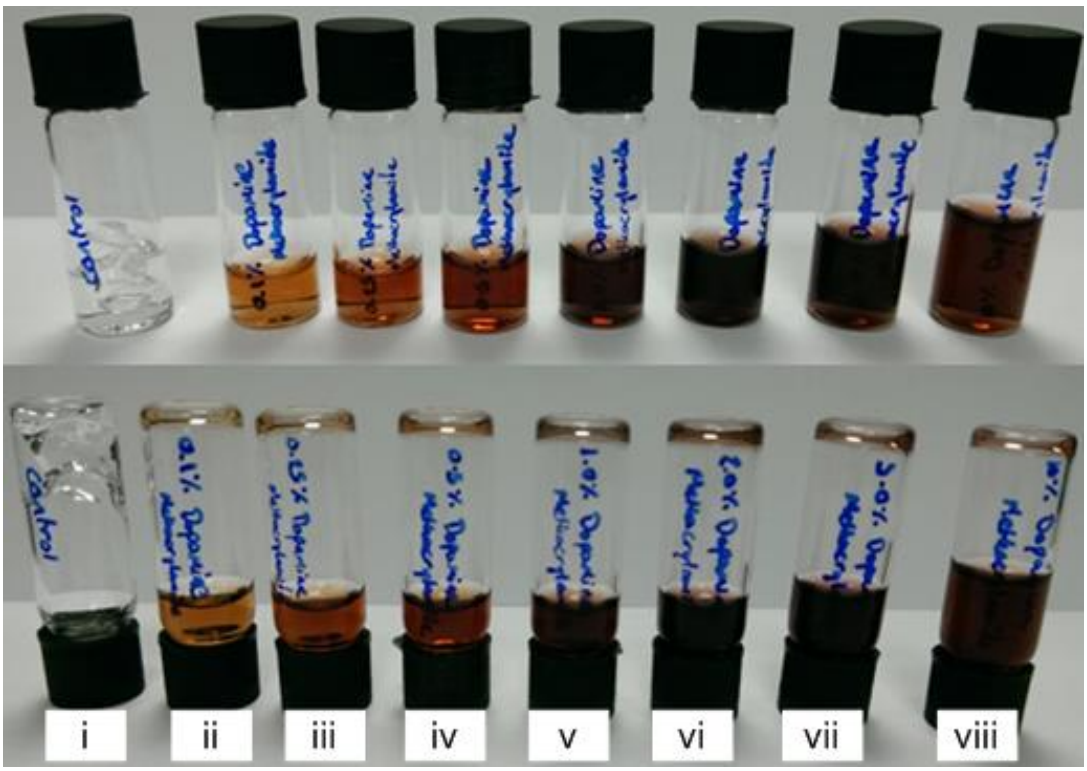


Figure 20: Sample gels prepared loaded with varying amounts of 47**. The percentage loadings of** 45 **in each sample are: i) 0.0 %; ii) 0.1 %; iii) 0.25 %; iv) 0.5 %; v) 1.0 %; vi) 2.0 %; vii) 5.0 %; viii) 10.0 %. Upon inversion of the vials it can be seen that samples ii through viii have not formed a gel, whereas polymerisation has occurred in sample i.**

Following the synthesis of **45**, it was decided to incorporate it into small scale acrylamide gels at a range of percentage loadings (0 % - 10 %) to determine the greatest amount of **45** that could be added without disrupting the polymerisation. To achieve this, a solution of **45** in acetonitrile was added to a mixture of tris(hydroxymethyl)aminomethane (Tris), polyacrylamide, and water, before the addition of the polymerisation initiators. These mixtures were left for approximately 30 minutes for the polymerisation to occur, however it was found that the solutions loaded with **45** did not undergo polymerisation (Figure 20).

It was suggested that the acetonitrile used in the solution of **45** could have interfered with the polymerisation, so it was decided to determine if another solvent could be used in the solution that would not interfere with the formation of the gels. It was found that **45** dissolved readily in dimethylformamide (DMF), so further sample gels were prepared following the same method as before but instead using a solution of **45** in DMF. Unfortunately, it was found that gel sample prepared with this new solution also did not set (Figure 21).

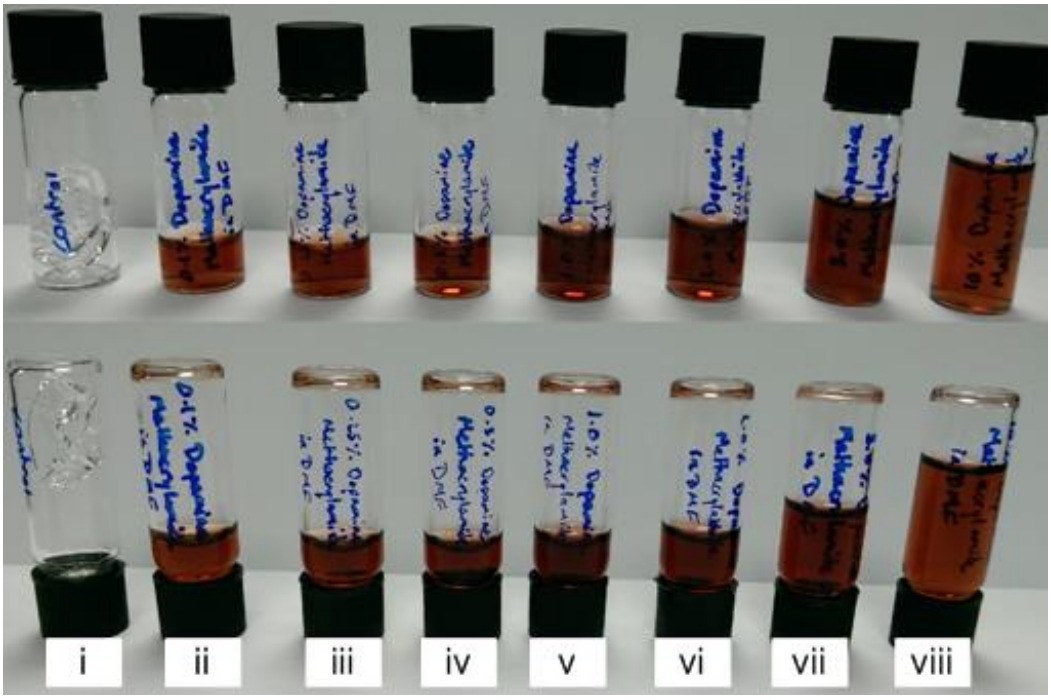


Figure 21: Sample gels prepared incorporating a solution of 45 **in dimethylformamide. The percentage loadings of** 47 **in each sample are: : i) 0.0 %; ii) 0.1 %; iii) 0.25 %; iv) 0.5 %; v) 1.0 %; vi) 2.0 %; vii) 5.0 %; viii) 10.0 %. When turning the vials it can be seen that sample i has successfully formed a gel, whereas no polymerisation has occurred for samples ii through viii.**

With both solutions of **45** leading to the same result, it was decided to prepare sample gels with a range of volumes of DMF to investigate whether or not it was the solvent or **45** disrupting the polymerisation. These gels were prepared as before but with DMF added instead of a solution of **45** (Figure 22). It can be seen that with 0.01 mL of DMF added, a gel

forms as readily as it does in the control sample. Increasing the volume of DMF to 0.10 mL shows some polymerisation but also a portion of the sample remains as a liquid. For the higher volumes of DMF (0.50 mL and 1.00 mL), whilst the solution exhibits turbidity, a gel is no longer formed, possibly due to the greater dilution of the acrylamide and other gel components.

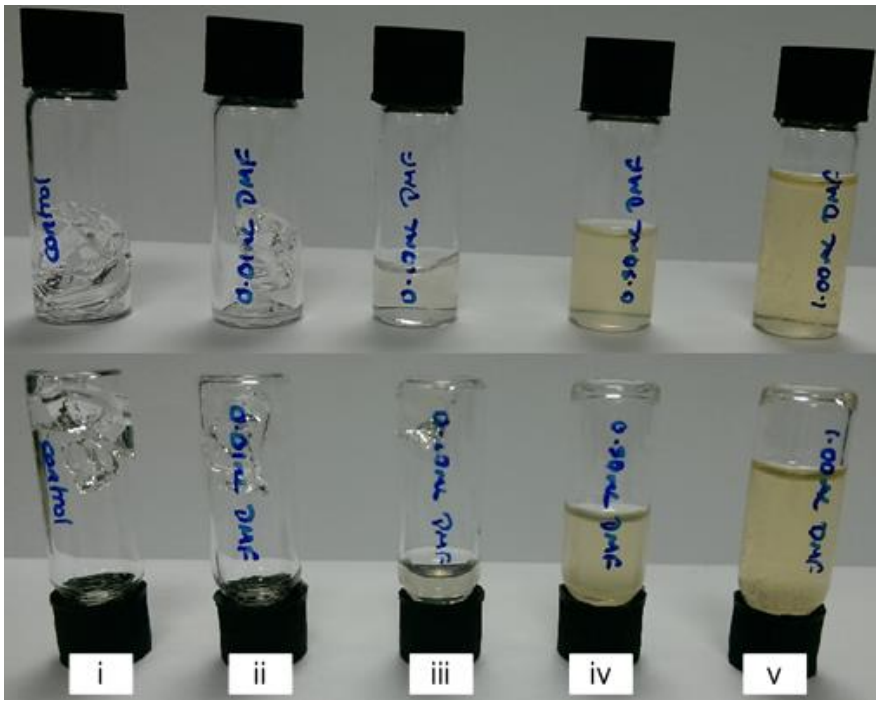


Figure 22: Sample gels prepared with added dimethylformamide. The amounts of dimethylformamide added to each sample are: i) 0.00 mL; ii) 0.01 mL; iii) 0.10 mL; iv) 0.50 mL; v) 1.00 mL. Polymerisation can be observed in samples i, ii, and iii, however samples iv and v have not formed gels.

The successful formation of a gel by the samples with small volumes of DMF (Figure 22) when no polymerisation is observed in samples with a small volume of the solution of **45** in DMF (Figure 21) suggests that **45** itself prevents polymerisation in the samples, rather than an the solvent. In light of these results, it was decided to attempt a different method for identifying labelled proteins in the electrophoresis gels. Stains and dyes are frequently used in electrophoresis to visualise protein samples, so it was decided to investigate if a selective stain for boronic acids could be used to visualise the labelled proteins. Curcumin and Alizarin Red S (ARS) were identified as promising compounds as both have been shown to detect compounds containing boronic acids.[120, 121]

Curcumin and ARS were added to samples of both labelled and unlabelled HSA to determine if the stains could be used to selectively identify labelled proteins. Both of the stains were added at a range of concentrations (curcumin: 0 - 0.27 M, ARS: 0 - 0.29 M) to investigate the amount required to stain the labelled samples while still allowing differentiation between labelled and

unlabelled proteins. The HSA samples mixed with the stains were then analysed by electrophoresis using an unmodified polyacrylamide gel.

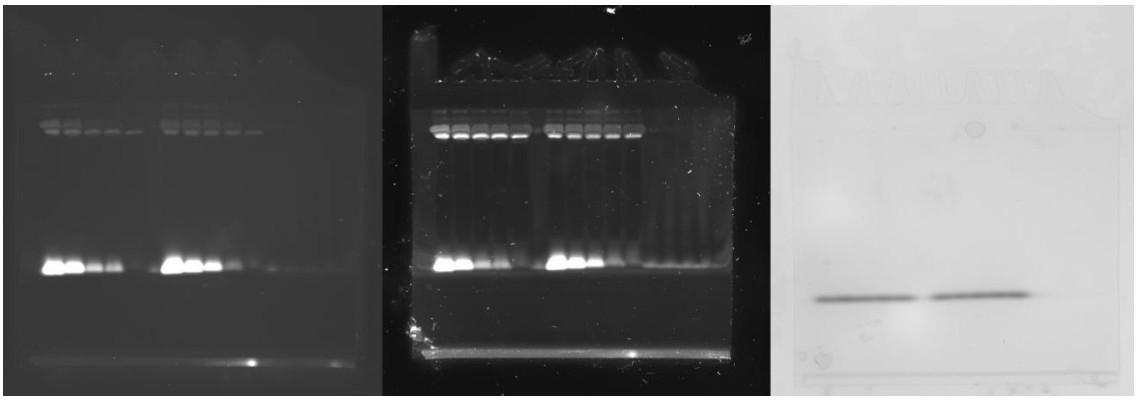


Figure 23: Electrophoresis gel of human serum albumin and labelled human serum albumin samples stained with curcumin illuminated with blue light (left), UV light (centre), and white light (right). From left to right the samples shown are: i - v) human serum albumin; vi - x) human serum albumin labelled with 39**. From left to right the concentrations of curcumin added are: i) 0.27 mM; ii) 0.14 mM; iii) 0.07 mM; iv) 0.03 mM; v) 0 mM; vi) 0.27 mM; vii) 0.14 mM; viii) 0.07 mM; ix) 0.03 mM; x) 0 mM.**

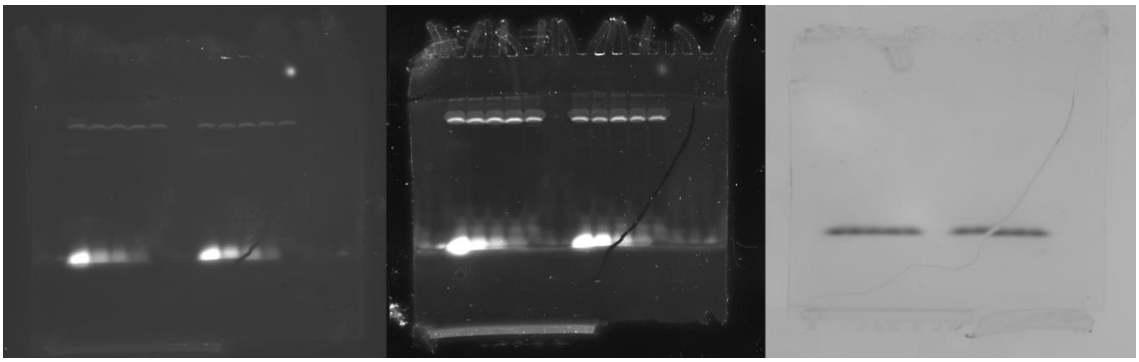


Figure 24: Electrophoresis gel of human serum albumin and labelled human serum albumin samples stained with Alizarin Red S illuminated with blue light (left), UV light (centre), and white light (right). From left to right the samples shown are: i - v) human serum albumin; vi - x) human serum albumin labelled with 39**. From left to right the concentrations of Alizarin Red S added are: i) 0.29 mM; ii) 0.15 mM; iii) 0.07 mM; iv) 0.03 mM; v) 0 mM; vi) 0.29 mM; vii) 0.15 mM; viii) 0.07 mM; ix) 0.03 mM; x) 0 mM.**

From Figure 23 and Figure 24, it can be seen that both curcumin and ARS do not selectively stain the labelled samples as theorised. It has been shown that curcumin interacts with HSA, explaining why curcumin does not selectively stain the labelled samples.[122] Whereas for ARS, this was suggested to be a result of the compound's charge. As the proteins are also charged, they could form an electrostatic interaction with ARS, which would exhibit no selectivity. It can be seen in Figure 24 that the samples stained with ARS are identical across all concentrations regardless of the presence of the boronic acid, **39**, which supports this suggestion.

Despite this lack of selectivity, in Figure 23 there appears to be a difference in the intensity of the bands under both blue and UV light. This difference is clearest at lower concentrations of curcumin (0.03 M and 0.07 M), whilst at higher concentrations there appears to be very little

difference. These data suggest that staining with curcumin holds some promise for identifying labelled proteins in gel electrophoresis, however the results are not as conclusive as initially hoped.

3.5 Conclusions

Through mass spectrum analysis, it has been shown the photoaffinity probe, **39**, can successfully label both HSA and MIF. However, it appeared that the labelling procedure resulted in the modification of casein leading to complications in confirming if the protein had been labelled. This shows that whilst **39** can label glycated proteins the method for its use proposed here is not suitable for all proteins. Control experiments highlighted that **39** is selective for glycated proteins and demonstrated that both the targeting and binding units were necessary for the labelling to occur.

From initial experiments with HSA it was determined that a 5:1 molar ratio of probe:protein was most suitable. It became clear with only a small number of experiments that using a greater amount of probe than this would cause the signal strength in the mass spectrometry to suffer and make analysis more difficult. This was suggested to be a result of the wider range of possible products from the labelling. While no ratio smaller than 5:1 was used, however, it was decided that the addition of a smaller amount of probe would not produce a change in mass significant enough to provide conclusive evidence that the labelling had been successful.

Investigation into other possible methods for analysing the labelled proteins focused on using gel electrophoresis. Drawing inspiration from to boronate affinity gel electrophoresis,[72] it was decided to attempt to make a diol containing gel that would selectively interact with the boronic acid labelled proteins to separate labelled and unlabelled proteins. However, it was found that diol selected, *N*-[2-(3,4-dihydroxy-phenyl)-ethyl]acrylamide, prevented polymerisation so no gel was successfully formed.

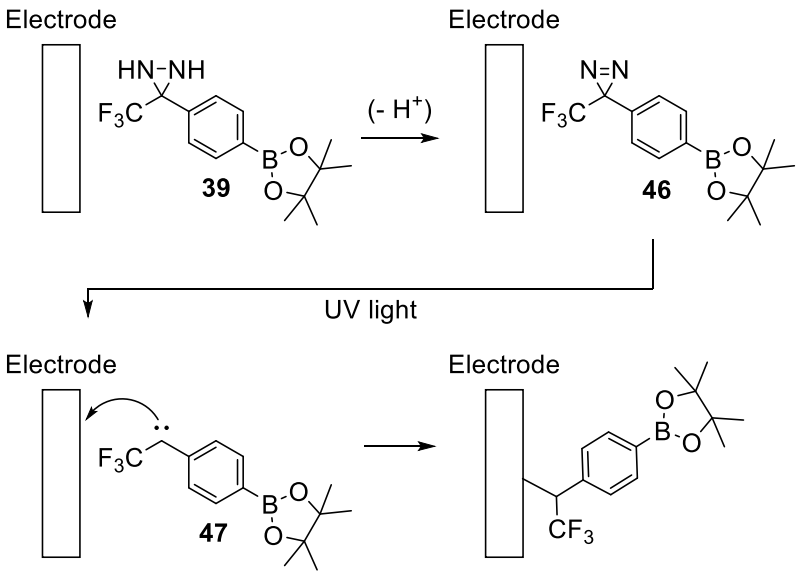
Following this it was decided to instead try a different method and use a compound to stain the labelled proteins on a non-functionalised polyacrylamide gel. Two stains were investigated, curcumin and Alizarin Red S. Curcumin did not exhibit the desired selectivity, however a difference could be seen in the intensity of the samples when observed under blue or UV light. ARS, however, was found to be unsuitable as it stained labelled and unlabelled proteins equally, suggested to be a result of it being charged. The results with curcumin did not give the immediate indicator of labelled proteins desired but do offer a suggestion of promise for this method.

Chapter 4: Electrochemistry

Electrochemical sensors are commonly used in a diverse range of areas, including in industrial processes, environmental studies, observing biological process, and medical measurements. These sensors tend to have lower precision and higher detection limits than other methods of spectroscopy, however they are simpler to set up and use and require less maintenance.[123] Additionally, they are convenient as they allow for real time analysis of samples and can detect compounds with the need for pre-treating samples.[124]

Electrochemical biosensors for glycoproteins[124] and, more generally, saccharides[125] have been an area of interest due to the biological importance of these molecules. As previously discussed, many biomarkers for diseases are glycoproteins; carcinoembryonic antigen and α-fetoprotein are recognised are markers for cancer, while glycated haemoglobin can be used to detect hyperglycemia.[124] The sensors used for such molecules are typically created by modifying the surface of an electrode with a compound that will bind to the target molecule. The typical receptors compounds used include: antibodies, lectins, and boronic acids.[124] Sensors modified with boronic acids have been shown to be effective in the detection of both sialic acid[126] and sugars.[127]

It was suggested that **39** could be used to fabricate a functionalised electrode which could then be used in an electrochemical sensor . This was performed to search for a faster and simpler method for identifying glycated proteins than analysing with mass spectrometry.



Scheme 25: Proposed method of using 39 **to functionalise an electrode. Following the oxidation of** 39 **into** 46**, illumination with ultraviolet light generates a carbene (**47**) at the electrode. Due to its reactivity, it is suggested the carbene will then form a covalent linkage with the surface of the electrode.**

Scheme 25 shows the proposed method for functionalising an electrode. By electrochemically oxidising **39** into the diazirine, **46**, and then irradiating the cell with UV light, it was thought it would be possible to generate a carbene directly at the electrode. This would then bind with the surface of the electrode in the same manner as shown when this molecule was used to label proteins. Within this chapter, the electrochemistry of the photoaffinity probe is discussed and the effect of treating an electrode by this method is investigated.

4.1 Electrochemical Behaviour of the Photoaffinity Probe

It was decided to investigate the electrochemical behaviour of **39** before attempting to functionalise the surface of an electrode. To achieve this, cyclic voltammograms were obtained of **39** to identify the oxidation of the diaziridine.

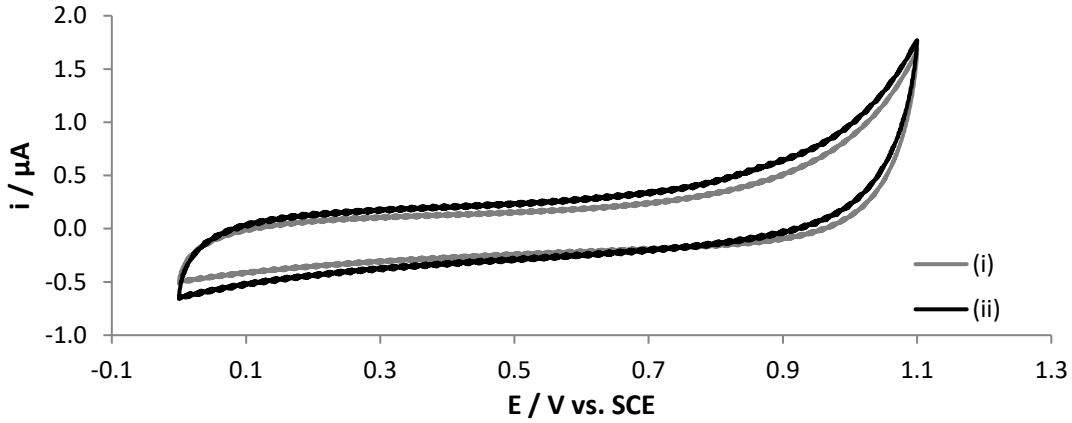


Figure 25: Cyclic voltammogram of (i) no additive, and (ii) 3.1 mg of 39**, in 10.0 mL of 0.1 M KNO_3 solution, scan rate 100 mVs^{-1}, using a glassy carbon working electrode and a platinum wire counter electrode.**

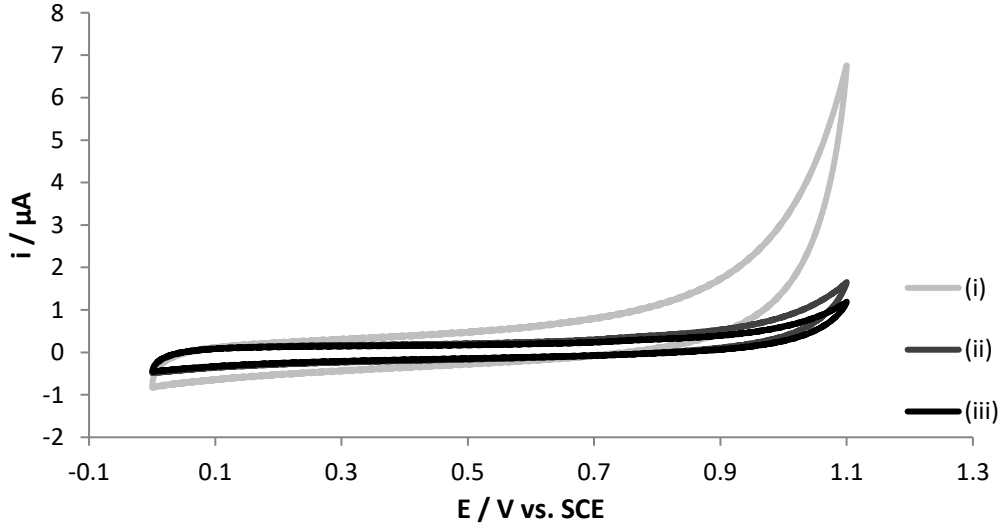


Figure 26: Cyclic voltammogram of (i) no additive, (ii) 1.7 mg in the dark, and (iii) 1.7 mg of 39**, in 10.0 mL of phosphate buffer (0.1 M) pH 7.2, scan rate 100 mVs^{-1}, using a glassy carbon working electrode and a platinum wire counter electrode.**

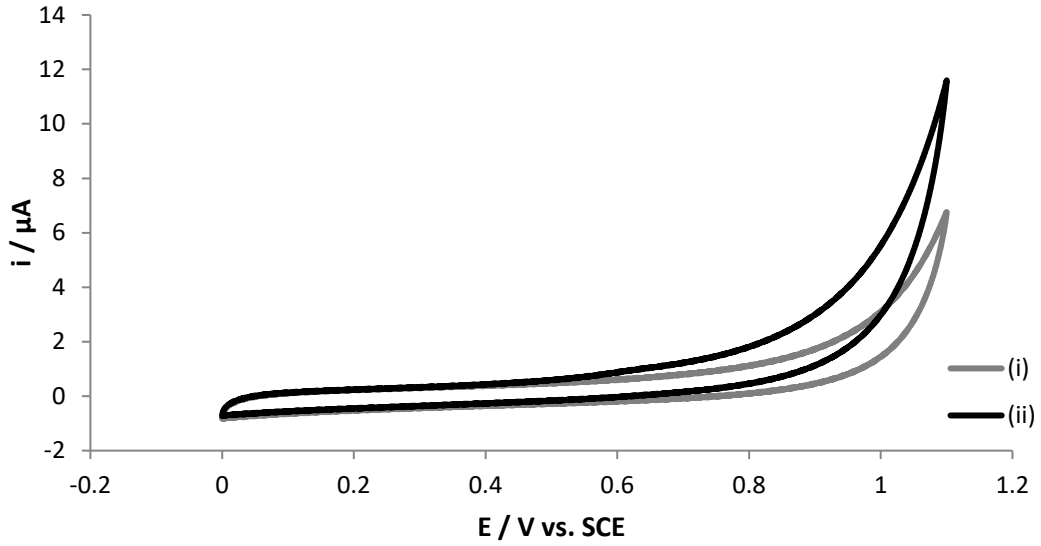


Figure 27: Cyclic voltammogram of (i) no additive, (ii) 1.7 mg of 39**, in 10.0 mL of phosphate buffer (0.1 M) pH 10.4, scan rate 100 mVs^{-1}, using a glassy carbon working electrode and a platinum wire counter electrode.**

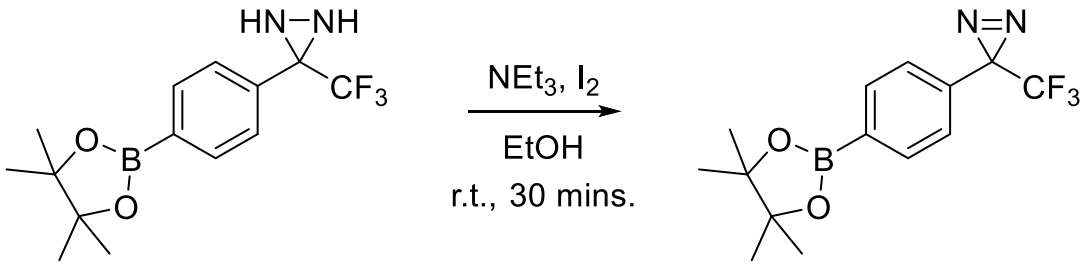
The cyclic voltammetry of **39** was initially performed in a 0.1 M potassium nitrite solution. However, in Figure 25 it can be seen that there is no marked difference between the behaviour of the electrolyte alone and a solution containing **39**. This shows that the expected oxidation to form a diazirine has not occurred. Following this, the experiment was reattempted using a potassium phosphate buffer solution at pH 7.2 as the electrolyte instead. It was also suggested that ambient UV light may have caused problems in observing the electrochemical events due to the sensitivity of the diazirine. To investigate this possibility, the experiment was performed in both light and dark. Figure 26 shows no difference between the experiments performed in the light and dark and that in both cases no oxidation of **39** is observed. As a result, it was decided to try the phosphate buffer again using a higher pH (10.4) in case this factor affected the progress of the oxidation. As seen in Figure 27, increasing the pH of the electrolyte did not result in any visible redox event in the cyclic voltammogram.

These results suggest that **39** cannot be oxidised through this method. A literature search confirmed that there is little precedent for electrochemical studies of diaziridines, although diazirines have been previously studied in this capacity.[128] The authors of that study show that electrochemical reduction of a diazirine is possible, however the nature of this reduction depended on the compound used and the presence of a proton source. While it was demonstrated that chlorinated diazirines would reduce irreversibly to form nitriles, under the same conditions diazirines without good leaving groups were shown to be reversibly reduced in to a radical monoanion. The addition of a proton source in the latter case altered the behaviour and instead the irreversible formation of the corresponding diaziridine was

observed.[128] This supports the conclusion drawn from the cyclic voltammograms obtained here (Figures 25, 26, and 27), that a diazirine cannot be produced by electrochemically oxidising a diaziridine using the methods described.

In light of this conclusion, it was decided to investigate the electrochemical behaviour of the diazirine. It was suggested that this would be of interest because, while the -CF_3 moiety could act as a leaving group, it might be possible to achieve the aim of generating a diazirine at the electrode by performing a cyclic voltammogram on the diazirine itself. If the electrochemical behaviour of **46** matched that reported for the non-halogenated diazirines,[128] the diazirine would be regenerated during the back oxidation.

In order to obtain cyclic voltammograms of **46**, the compound first needed to be synthesised and isolated as earlier uses had generated **46**, and subsequently the carbene, *in situ*. The synthesis followed the same procedure as previously used, oxidising **39** with a combination of triethylamine and iodine (Scheme 26). The reaction and subsequent work-up were performed in the dark to avoid deterioration of **46**, which was successfully isolated with a good yield (92 %).



39 **46**

Scheme 26: Oxidation of 39 **to form a diazirine,** 46**.**

Following the synthesis of **46**, cyclic voltammograms were obtained using a solution of tetrabutylammonium hexafluorophosphate in acetone as the electrolyte. This was initially performed with only a small amount of **46** (3.0 mg) and as seen in Figure 28, a reduction event was visible. However, the back oxidation produced a much smaller peak. The amount of **46** was then increased to 6.0 mg and a further cyclic voltammogram obtained. As expected, Figure 28 shows that the reduction and subsequent oxidation events are more pronounced with a greater concentration of **46.** This makes the difference between the reduction and oxidation peaks clearer but confirms that the reduction is reversible. This result is in line with the previously reported data for the non-halogenated diazirines, suggesting that -CF_3 is not a sufficiently good leaving group to allow for the irreversible formation of a nitrile.[128]

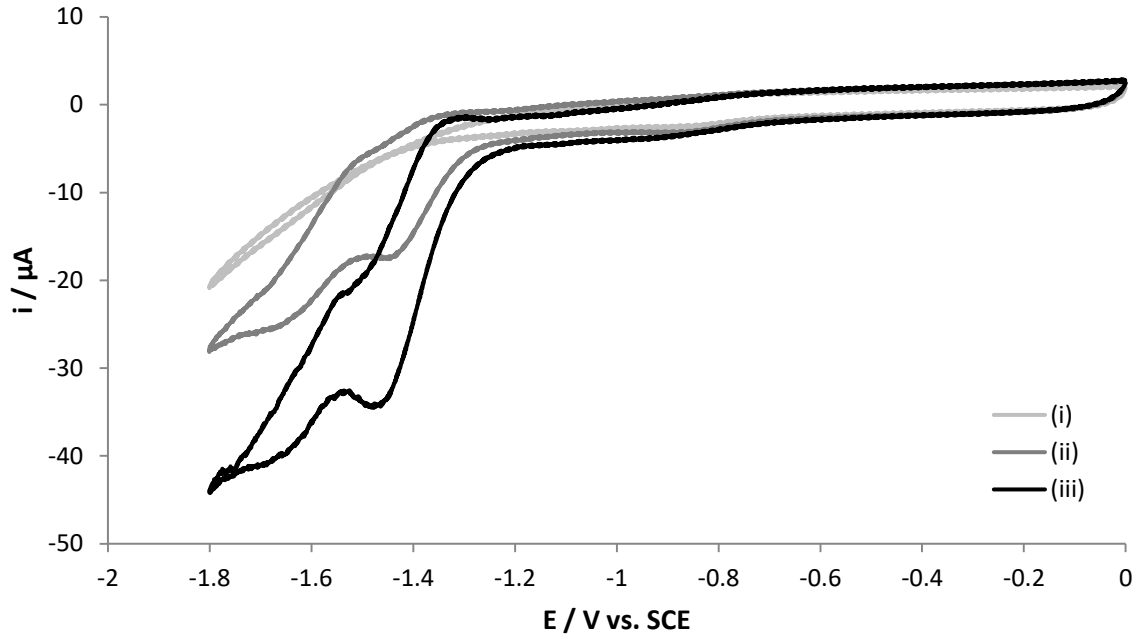


Figure 28: Cyclic voltammogram of (i) no additive, (ii) 3.0 mg of 46**, and (iii) 6.0 mg of** 46**, in 10.0 mL of a tetrabutylammonium hexafluorophosphate solution in acetonitrile (0.1 M), scan rate 100 mVs^{-1}, using a glassy carbon working electrode and a platinum wire counter electrode performed in the dark.**

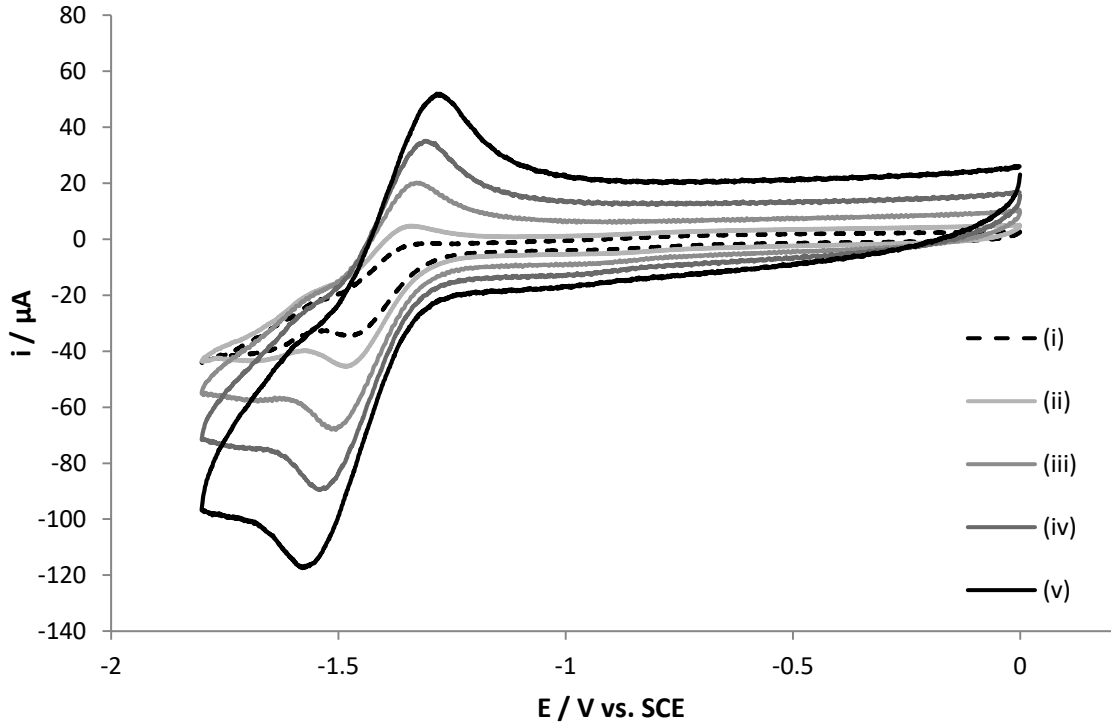


Figure 29: Cyclic voltammogram of 6.0 mg of 46 **in 10.0 mL of a tetrabutylammonium hexafluorophosphate solution in acetonitrile (0.1 M), scan rates: (i) 100 mVs^{-1}; (ii) 200 mVs^{-1}; (iii) 500 mVs^{-1}; (iv) 1000 mVs^{-1}; (v) 2000 mVs^{-1}, using a glassy carbon working electrode and a platinum wire counter electrode, performed in the dark.**

It was noted by the authors of the reported study that the reversibility of the reduction was enhanced for fast scan rates and low temperatures. The cyclic voltammograms were therefore

repeated using 6.0 mg of **46** at a range of scan rates. As reported in the literature, Figure 29 shows that for the slower scan rates (100 and 200 mVs^{-1}) the peak produced by the back oxidation is noticeably smaller than that of the reduction. However, from Figure 29 it can also been seen that there is an increase in the parity between the extent of the reduction and oxidation peaks as the scan rate increases from 100 mVs^{-1} up to 2000 mVs^{-1}. This provides further evidence that the process observed here matches the single electron reduction to form a radical monoanion proposed in the literature and that the following oxidation regenerates the diazirine.[128]

4.2 Surface Functionalisation

After demonstrating that the electrochemical reduction and back oxidation of **46** would form the diazrine at the electrode as desired, it was decided to investigate using **46** in place of **39** in the proposed method to functionalise an electrode (Scheme 25). For these experiments, it was suggested that a fluorine doped tin oxide (FTO) glass electrode would be a suitable target for functionalisation as examples of the use of these electrodes in similar work exist.[129]

To achieve this, an electrochemical cell with an FTO working electrode was used to reduce and re-oxidise **46** in the dark before irradiation with UV light. This electrode was subsequently immersed in ARS before being rinsed with deionised water, this followed a reported method used to visually identify boronic acids attached a surface through staining.[130] Figure 30 shows that the FTO electrode treated with **46** has only minor colouration following the ARS dip and is not fully stained as expected, which would suggest that surface has not been functionalised with the boronic acid to the extent hoped, if at all.

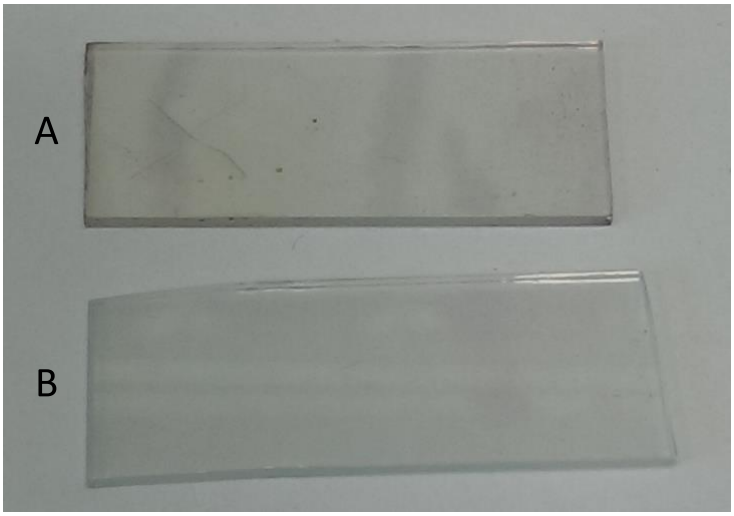


Figure 30: FTO electrode treated with 46 **after immersion in an Alizarin Red S solution (1.0 mM) for 60 seconds and rising with deionised water (A), beside an untreated FTO electrode (B) for comparison.**

As the staining result was not wholly conclusive, cyclic voltammograms of ARS were obtained using both the treated and untreated electrodes in order to obtain confirmation of the result of treating the electrode with **46**. It has been shown that, when ARS is bound to another compound, the potentials of peaks observed in the cyclic voltammogram are shifted.[131, 132] Figure 31 shows a clear shift in the potentials of the peaks between the treated and untreated electrodes, matching the reported cyclic voltammograms comparing the behaviour of complexed and free ARS.[131] This suggests that the boronic acid has successfully adhered to the surface of the FTO electrode. The apparently contrary observance of little visible staining following the ARS dip could indicate that only a small amount of the boronic acid has attached to the electrode. Alternatively, it could indicate that the staining process was not optimally tuned for this electrode and either the electrode required a longer time in the ARS solution or fewer rinses were necessary.

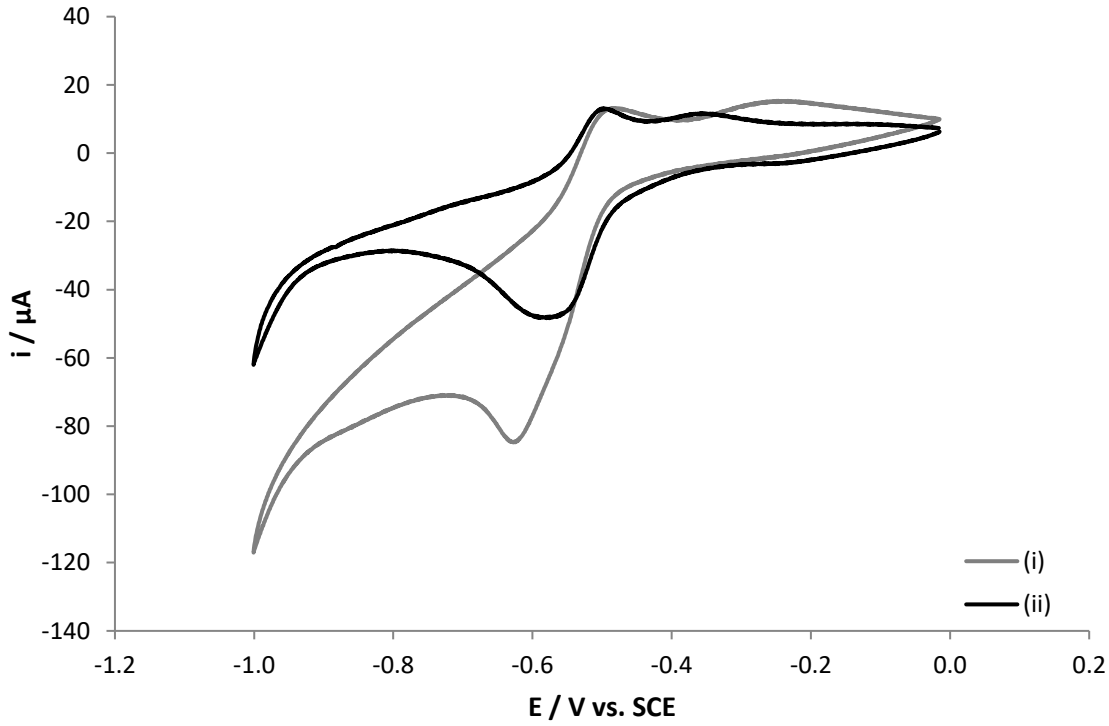


Figure 31: Cyclic voltammograms of Alizarin Red S in phosphate buffer (0.1 M) at pH 7.2, scan rate 100 mVs^{-1}, using (i) an untreated FTO electrode, or (ii) an FTO electrode treated with 46, with a platinum wire counter electrode.

4.3 Conclusions

The electrochemical behaviour of **39** has been investigated with the intent of generating a diazirine, and subsequently a carbene, at the electrode through oxidation of **39** to facilitate functionalisation of the surface. It was shown by the cyclic voltammograms of **39** obtained that, under the conditions used, neither the desired oxidation nor any other redox events occurred.

Without being able to form the diazirine, the suggested method was no longer viable and an alternative was investigated.

An electrochemical study of diazirines in the literature[128] showed that it might be possible to modify the proposed method to use the diazirine, **46**, in place of **39**. It was suggested that following the electrochemical reduction of **46** it would be possible to regenerate the diazirine at the electrode during the back oxidation.

Subsequently, **46** was successfully synthesised and isolated before cyclic voltammograms were obtained. These data indicated that **46** could be reversibly reduced within an electrochemical cell. This behaviour was then used to generate **46** at an FTO electrode, which was irradiated with UV light. The treated FTO electrode was then analysed with an ARS dip, which produced only a minor visible change when compared to an unchanged FTO electrode. Further confirmation of the result of the treatment was sought by obtaining cyclic voltammograms of ARS with both an untreated FTO electrode and the treated electrode. The results of these showed that for the treated FTO the potential of the redox events was shifted, matching previously reported data comparing the cyclic voltammograms of free and bound ARS.[131] This suggests that some of the boronic acid has been successfully attached to the surface, showing that the diazirine, **46**, has potential for application in surface functionalisation, possibly leading to electrochemical sensors; however, refinement of the method and further analysis of the surfaces produced is required.

Chapter 5: Conclusions

This thesis has detailed the work undertaken to design, synthesise, and investigate the uses of, a novel photoaffinity probe. This probe was designed to use a boronic acid (**30**) moiety to selectively target glycated proteins, exploiting the known and well documented affinity of boronic acids for sugar molecules.[64, 65, 133] Complications were met during the purification process, and it was ultimately decided to form the boronate ester with pinacol allowing for easier isolation of the target compound. Following purification, the ester was not converted back as it was proposed that **46** would be just as able to target a glycated protein as **30** due to the increased favourability of boronate esters formed with sugar molecules.[88] It was decided to store the probe as the diaziridine analogue (**39**), due to the light sensitivity of the diazirine (**46**).

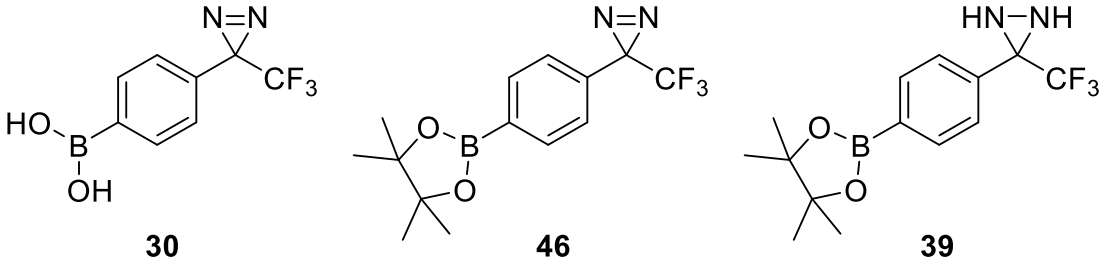


Figure 32: The initial (30**) and final (**46**) designs of the photoaffinity probe, alongside the diaziridine analogue (**39**) which the probe was stored as to its relative stability.**

The initial synthetic route followed a previously described method for the generation of a similar compound,[85] which gave, at best, moderate overall yield of 59 %. This left some room for improvement and, with a high variation in the yields observed for some of the steps in the initial route, it was decided to investigate possible optimisations. This led to a number of different possible routes incorporating Suzuki-Miyaura coupling reactions. The alternative routes devised gave overall yields of 53 % and 39 %, both lower than the initial path. However, the complexity of these routes was also lower as they avoided air and moisture sensitive intermediates. Additionally, the conditions used in these routes had not been tailored to the compounds. Considering these points, the more efficient of the two alternative routes would appear to offer another good option for the synthesis of **39**, and ultimately **46**.

Following the investigations into the synthesis, the ability of the photoaffinity probe to label glycated proteins was demonstrated. Through mass spectrometry, it has been shown that the probe was able to successfully label the proteins HSA and MIF. In these experiments, the labelled protein samples showed an increase in mass over their unlabelled counterparts that

was in line with the value expected. A third protein, casein, was also investigated, however the result obtained showed an unexpected loss of mass. From this the conclusion drawn was that the labelling procedure had inadvertently damaged the protein structure, possibly due to the exposure to UV light. This showed that, while the probe did work as intended, the method detailed in this thesis is not suitable for use with all proteins.

Two different ratios of probe to protein were initially trialled with HSA, revealing that at a high loading the strength of the mass spectrum recorded was reduced. This was surmised to be a result of a wider range of labelling products leading to a dilution of the signal. It was demonstrated that a 5:1 ratio of probe to protein gave both a clear increase in mass and retained a good signal strength.

Having shown that the labelling was successful, control experiments were performed to show that the labelling was a result of the combination of the targeting effect of the boronic acid and the binding ability of the carbene generated from the diazirine. These experiments used molecules that mirrored **39** but with each lacking one of the two integral groups. When either functionality was missing, no increase was observed in protein mass following the labelling protocol. These results show that **39** acts as proposed, with the boronate ester moiety targeting by reversibly binding with sugar molecules on the protein surface before the diazirine is transformed into the carbene, which binds irreversibly to permanently label the target. Additionally, the selectivity of **39** was demonstrated by comparing the result of performing the labelling procedure on both glycated and non-glycated MIF; mass spectrum analysis showed the expected increase for the glycated protein and no change for the latter.

While mass spectrometry allowed for detailed and accurate analysis of the labelled samples, finding a simpler and faster method for identifying labelled proteins was desirable. Efforts towards this initially focused on creating a modified electrophoresis gel, using the same theory as boronate affinity electrophoresis, albeit reversed.[72, 73] Attempts to produce a polyacrylamide gel incorporating a diol were unsuccessful as the diol used prevented the polymerisation of the gel. Another approach investigated was the use of stains to selectively highlight labelled samples. Curcumin and Alizarin Red S were selected to trial as compounds commonly used in boron detection.[120, 121] ARS displayed no visible selectivity for labelled samples, concluded to be a result of it being a charged compound. Curcumin, however, showed a slight difference between labelled and unlabelled samples under UV light and blue light. This indicates that this method of staining could be employed to more quickly and easier identify the presence of labelled proteins.

The electrochemical behaviour of the diaziridine, **39**, has been detailed here, showing that, under the conditions tested, **39** cannot be electrochemically oxidised to produce its diazirine analogue. Cyclic voltammograms obtained of **46**, however, display a distinct reversible reduction. By comparing these results with previously reported data in the literature,[128] it can be seen that the values obtained here match those for a single electron reduction leading to the reversible formation of a radial monoanion of the diazirine. This provides some insight into the electrochemistry of **46**, showing that it behaves as other previously reported diazirines.

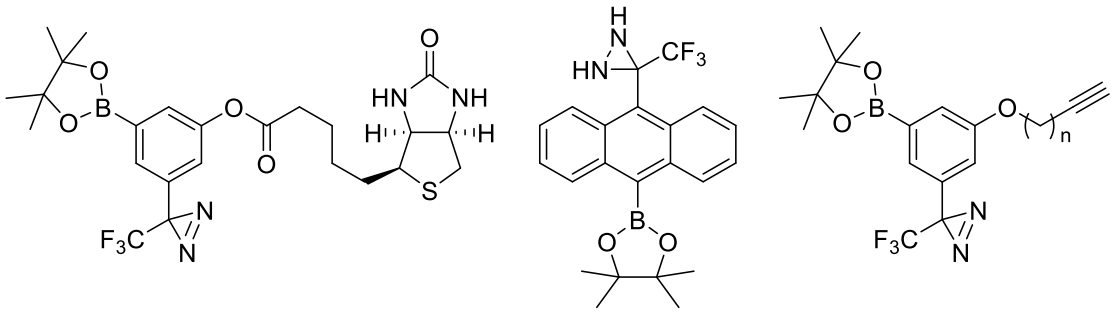
After obtaining these cyclic voltammograms, **46** was employed in functionalising the surface of an electrode. An FTO electrode following treatment with **46**, was dipped in ARS, however the expected staining was not observed and only minor colouration could be seen. Electrochemical tests were then carried out for further clarification. The treated FTO showed a shift in the potential of the peaks seen in a cyclic voltammogram of ARS as compared to an untreated electrode. These changes match literature values comparing the electrochemistry of free and bound ARS.[131]

This thesis shows that boronic acids can be employed in the development of photoaffinity probes for the selective labelling of glycated proteins. The protocol used herein is a simple and effective method for the treatment of samples, which can then give a clear indication of glycation through standard protein analysis techniques. An alternative possible use has been demonstrated in a straightforward method for surface functionalisation.

Chapter 6: Future Work

The possibility of using a boronic acid as the targeting part of a photoaffinity probe for glycated proteins and the alternative use of such a probe in surface functionalisation have been demonstrated by the work presented in this thesis. Yet, there exists a wide scope for further investigation into this area and the possible uses of such molecules.

A preliminary optimisation of the synthesis of **39** has been detailed in this thesis. However, as discussed previously there is room for more work on this. Considering just the synthetic routes presented in this thesis, there is potential for improvement in a number of the steps, particularly in the routes that included a Suzuki-Miyaura coupling. As previously mentioned, the conditions used for these reactions were the same as those for the initial synthetic route, which followed the previously reported method.[85] It is entirely possible that the incorporation of a boronate ester into the starting materials would lead to a different set of optimal conditions. Even by studying alternative solvent systems and temperatures alone, a more efficient method for performing these synthetic steps could be revealed.



48 **49** **50**

Figure 33: Possible designs for other boronic acid-based photoaffinity probes. A probe containing biotin as a pre-attached reporter (48**), an anthracene based probe (**49**) proposed a possible fluorescent alternative, and a probe with an alkyne handle for reporter attachment (**50**).**

Beyond investigations into the synthesis of **39**, possible modifications to the structure of the probe (Figure 33) could be considered, now that it has been shown that such probes can successfully fulfil their purpose in labelling glycated proteins. It would be of interest to develop a probe with a reporter group to allow for easier detection, a fluorescent analogue for example, or by including functionality like an alkyne or azide handle for easy attachment of a reporter after labelling. This would give alternative methods for analysis of the labelled samples that would not rely on mass spectrometry. A fluorescent probe (**49**) would be easily detectible in an electrophoresis gel, while attaching biotin to the probe (**48**) would enable detection through immunohistochemistry or isolation with affinity chromatography. The

inclusion of a functionalisation handle (**50**) would allow for the attachment of a reporter group at a later stage, while this would require a further synthetic step it would allow for greater versatility.

Alternatively, further work into the creation of functionalised gel could lead to easier analytical methods without the need to synthesise a new probe. While the initial investigations into this method were unsuccessful, a broad study of alternative diols for could identify a molecule that can be incorporated into the gel without hampering polymerisation. This would then allow for an investigation into the efficacy of these types of gel in separating proteins with a boronic acid moiety attached and those that are unlabelled.

On the topic of sample analysis, the work reported in this thesis on the use of stains to identify labelled samples provides an opportunity for further investigation. The results obtained with HSA and curcumin indicated the potential of this method, although the lack of selective staining thought to be the result of an HSA-curcumin interaction led to these results being less clear than desired. Expanding this study to investigate labelled samples of other proteins and also other possible stains would provide evidence to clarify the suitability of this method. Additionally, the fluorescence of curcumin combined with both labelled and unlabelled HSA, and other proteins, could be quantified through fluorimetry. This analysis would give greater insight into the detection of labelled protein samples with curcumin.

The cyclic voltammograms of ARS obtained using a treated electrode suggest the successful functionalisation of the surface, despite the lack of visible staining following an ARS dip. For greater detail, the resulting surface could be further analysed using techniques such as Raman spectroscopy or atomic force microscopy.

The potential uses of an electrode treated with **46** provides another an area of study to explore. The functionalisation was initially investigated due to the abundance of electrochemical sensors for biomolecules and it would be of interest to study the potential for using an electrode functionalised through this method in a sensor.

Chapter 7: Experimental

7.1 General Procedures

All solvents and reagents were reagent grade unless otherwise stated and were purchased from Acros Organics, Fisher Scientific UK, Frontier Scientific Europe Ltd, or Sigma-Aldrich Company Ltd and were used without need for further purification. Human serum albumin and casein samples were purchased from Sigma-Aldrich Company Ltd and recombinant protein samples were synthesised by Dr Omar Kassaar Munir of the Biology and Biochemistry department. Anhydrous solvents were obtained from an Innovative Technology solvent purification system. Glassware used in reactions involving moisture sensitive compounds was dried in an oven at 200°C overnight before use.

Silica gel column chromatography was performed using Davisil Grade 633 silica gel with a 60Å pore size purchased from Sigma-Aldrich Company Ltd.

Nuclear magnetic resonance spectra were recorded in $CDCl_3$, d_3-MeOD, DMSO-d_6, or D_2O or a combination of these solvents, the solvent used is noted with the data reported below. ^{1}H NMR spectra were taken with a Bruker Avance 250, Bruker Avance 300, or Agilent 500 spectrometer with operational frequencies of 250.13 MHz, 300.22 MHz, and 500.06 MHz respectively. ^{1}H-^{13}C NMR spectra and ^{11}B NMR spectra were recorded using a Bruker Avance 300 or Agilent 500 spectrometer at operational frequencies of 75.50 MHz and 125.74 MHz respectively for ^{13}C NMR spectra, and 96.32 MHz and 160.44 MHz for ^{11}B NMR spectra. ^{19}F NMR spectra were obtained using an Agilent 500 spectrometer with an operational frequency of 470.53 MHz. Chemical shifts (δ) in ^{1}H and #^{1}H#-^{13}C NMR spectra are given in parts per million (ppm) and reported relative to the residual solvent peak. In ^{11}B NMR spectra, chemical shifts (δ) are expressed in parts per million (ppm) and reported relative to boron trifluoride diethyl etherate as an external standard. In ^{19}F NMR spectra, chemical shifts (δ) are given in parts per million (ppm) and reported relative to trichlorofluoromethane as an external standard.

The procedure for labelling protein samples was performed in a small sample vial, a solution of the target protein (2.0 mg/mL) was mixed with a solution of the photoaffinity probe (1.0 mg/mL), quantities for specific proteins are listed later in this chapter. This vial was sealed whilst the mixture was kept in the dark. To expose the sample to ultraviolet light, the lid was removed and the open vial placed underneath an ultraviolet lamp used to visualise thin layer chromatography plates.

NMR Spectra were analysed using Bruker TopSpin 3.5 pl 5. The generic assignments and multiplicities of peaks in NMR spectra are: singlet (s), doublet (d), doublet of doublet (dd), triplet (t), triplet of doublet (td), quartet (q), unresolved multiplet (m), broad (br) and aryl (Ar).

The mass spectra of small molecules were recorded either in house with an automated Bruker Daltonics MicrOTOF mass spectrometer, or sent to the EPSRC UK National Mass Spectrometry Facility at Swansea University. Samples were ionised with electrospray ionisation in either positive or negative mode. Intact protein mass spectra were obtained by Mr Mervyn Lewis on an Agilent ESI-QTOF mass spectrometer.

Polyacrylamide gels for electrophoresis were made before use with a protocol used by the Department of Biology and Biochemistry. A 15 % separating gel was made by combining water (2.313 mL), a solution of tris(hydroxymethyl)aminomethane in water (1.5 M, 1.563 mL), and a solution of acrylamide in water (5.6 M, 2.338 mL). To this mixture *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (2.813 µL, 18.76 µmol) and a solution of ammonium persulphate in water (0.4 M, 31.25 µL) were then added. This mixture was then loaded into a cassette and left for 30 minutes until polymerisation was complete. Once the separating gel had set, a stacking gel was made by combining water (2 mL), a solution of tris(hydroxymethyl)aminomethane in water (0.5 M, 0.7813 mL), and a solution of acrylamide in water (5.6 M, 0.3125 mL). To this mixture *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (3.125 µL, µmol) and a solution of ammonium persulphate (0.4 M, 15.625 µL) were then added. This mixture was then loaded into the cassette on top of the stacking gel. A plastic comb was then inserted into the top of cassette to form wells and then the gel left to set for 30 minutes. The gels were used immediately after being made.

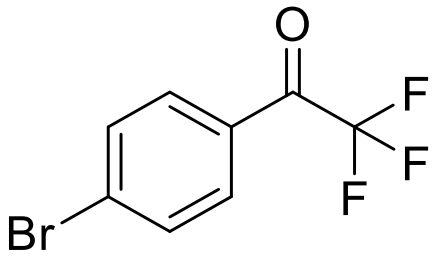
Functionalised gels were made using a small scale version of the protocol for making a 15 % separating gel. Water (0.370 mL), a solution of tris(hydroxymethyl)aminomethance in water (1.5 M, 0.250 mL), and a solution of acrylamide in water (5.6 M, 2.338 mL) were combined. To this mixture a solution of *N*-(3,4-dihydroxyphenethyl)methacrylamide in dimethylformamide or acetonitrile (0.45 M, 0.01 mL - 1.0 mL, 1.00 mg - 100 mg, 4.52 µmol - 452 µmol). To this mixture *N*,*N*,*N*′,*N*′-tetramethylethylenediamine (0.45 µL, 3.00 µmol) and a solution of ammonium persulphate in water (0.4 M, 5.00 µL) were then added. The mixture was then left to polymerise.

Cyclic voltammograms were obtained using a microAutoLab III with GPES software. The working electrode was either a glassy carbon electrode or a fluoride tin oxide electrode, with a

platinum wire counter electrode, and the reference electrode used was a saturated calomel electrode. Electrolyte, substrate, and electrodes were placed inside a large sample vial to make the electrochemical cell. For experiments in the dark, the cell was shielded from light by wrapping with aluminium foil. The electrolyte used was either: a solution of tetrabutylammonium hexafluorophosphate in acetonitrile (0.1 M); a phosphate buffer made a mixture of a monosodium phosphate solution (0.1 M) and a disodium phosphate solution (0.1 M) tuned to pH 7.2 or pH 10.4; or a potassium nitrate solution (0.1 M). The compound to be investigated would be dissolved into the cell electrolyte. The data was then analysed using Microsoft Excel.

7.2 Synthesis

4'-bromo-2,2,2-trifluoroacetophenone (32)[85]



Magnesium turnings (2.58 g, 106.0 mmol, 1.0 equiv.) were ground under nitrogen with a stirrer bar in a round bottomed flask to produce a dark grey powder. 1,4-dibromobenzene (25.00 g, 106.0 mmol, 1.0 equiv.) was then added to the powder, followed by anhydrous diethyl ether (250 mL). The mixture was then heated under reflux for 2 hours. After refluxing, the mixture was cooled to -78°C in a dry ice/acetone bath. Once cooled, ethyl trifluoroacetate (12.61 mL, 106.0 mmol, 1.0 equiv.) was added slowly and the resulting mixture was left to warm to room temperature overnight. After 16 hours the reaction was quenched with a saturated NH_4Cl solution. After separation the organic layer was washed with two portions of water (50 mL) and then with brine (50 mL) before drying with anhydrous magnesium sulphate. Removing the solvent under reduced pressure gave an orange oil, which was then purified by distillation (60 °C, 0.4 mbar) to afford the target compound as a clear oil (23.22 g, 91.8 mmol, 87 %).

^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 7.91 (d, *J* = 8.4 Hz, 2H, ArH), 7.69 (d, *J* = 8.4 Hz, 2H, ArH)

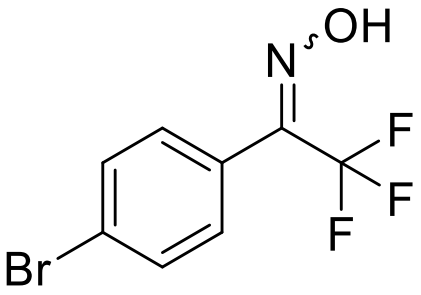
^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 179.87 (q, $^{2}J_{C\text{-}F}$ = 35.4 Hz), 133.32, 132.78, 131.56, 128.85, 116.71 (q, $^{1}J_{C\text{-}F}$ = 290.6 Hz)

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298K) δ_F (ppm): -71.69

HRMS (ESI-TOF) *m/z*: $[M+H]^+$ Calculated for $C_8H_5BrF_3O$ 252.9476; Found 252.9474

These NMR data match the reported results.[134]

1-(bromophenyl)-2,2,2-trifluoroethanone oxime (33)[85]



Method 1

4'-Bromo-2,2,2-trifluoroacetophenone (16.11 g, 63.7 mmol, 1.0 equiv.) was dissolved in a solution of ethanol (16 mL) and pyridine (32 mL). To the mixture, hydroxylamine hydrochloride (4.43 g, 63.7 mmol, 1.0 equiv.) was added. The reaction mixture was heated to 60°C and allowed to stir for 2 hours. After evaporating the solvent under reduced pressure, the residue was dissolved in diethyl ether (50 mL) and water (50 mL). The organic layer was separated and then washed with a further portion of water (50 mL) and brine (50 mL) before drying with anhydrous magnesium sulphate. Evaporation of the solvent yielded a yellow oil. When necessary, the product was purified by column chromatography to give the target compound (11.39 g, 60.2 mmol, 95 %). R_f = 0.57 (hexane/EtOAc, 3:1).

Method 2

4'-Bromo-2,2,2-trifluoroacetophenone (30.00 g, 118.5 mmol, 1.0 equiv.) was dissolved in a solution of ethanol (30 mL) and pyridine (60 mL). To the mixture, hydroxylamine hydrochloride (9.06 g, 130.4 mmol, 1.1 equiv.) was added. The reaction mixture was heated to 60°C and allowed to stir for 3 hours. After evaporating the solvent under reduced pressure, the residue was dissolved in diethyl ether (50 mL) and water (50 mL). The organic layer was separated and then washed with a further portion of water (50 mL) and brine (50 mL). A solution of sodium hydroxide (1.0 M, 50 mL) was then added to the organic phase. After separation, the aqueous phase was washed with three portions of diethyl ether (50 mL) before acidification with a solution of hydrochloric acid (1.0 M). Diethyl ether (50 mL) was then added to the aqueous mixture and the target compound extracted. The organic phase was then dried with anhydrous magnesium sulphate. Evaporation of the solvent yielded the target compound (19.32 g, 72.1 mmol, 61 %).

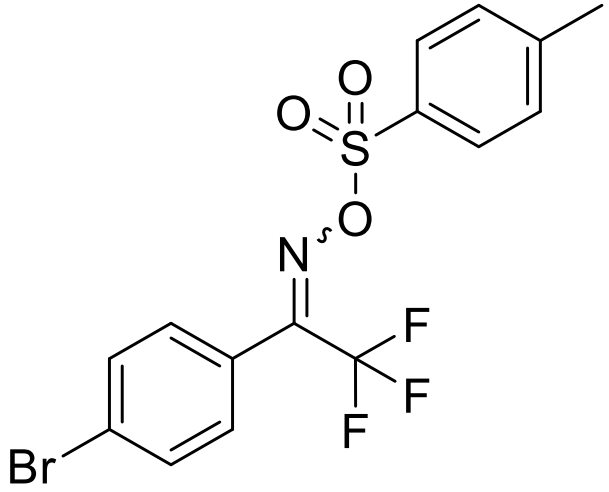
^{1}H NMR (Bruker AV300 @ 300.22 MHz, $CDCl_3$, 298K) δ_H(ppm): 10.18 (1H, br s, NOH), 7.69 (2H, d, *J* = 8.3 Hz, ArH), 7.50 (2H, d, *J* = 8.3 Hz, ArH)

^{13}C NMR (Bruker AV300 @ 75.50 MHz, CDCl3, 298K) δ_C(ppm): 146.64 (q, $^{2}J_{C\text{-}F}$ = 32.3 Hz), 132.00, 130.39, 129.96, 125.45, 120.53 (q, $^{1}J_{C\text{-}F}$ = 274.8 Hz)

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298K) δ_F (ppm): -78.41

^{1}H and ^{13}C NMR data match reported results.[135] A literature search yielded no previous report of ^{19}F NMR data.

1-(4-bromophenyl)-2,2,2-trifluoroethanone *O*-tosyl oxime (34)[85]



1-(Bromophenyl)-2,2,2-trifluoroethanone oxime (18.00 g, 67.2 mmol, 1.0 equiv.) was dissolved in dichloromethane (80 mL). Triethylamine (11.23 mL, 80.59 mmol, 1.2 equiv.) and 4-dimethylaminopyridine (0.82 g, 6.72 mmol, 0.1 equiv.) were added to the solution. The mixture was then cooled to 0°C. Then *p*-toluenesulfonyl chloride (14.72 g, 77.23 mmol, 1.15 equiv.) was added in portions. Following the addition, the reaction mixture was allowed to warm to room temperature before stirring for a further 3 hours. After evaporation of the solvent under reduced pressure, the residue was dissolved in diethyl ether (80 mL) and washed with two portions of water (40 mL) and one of brine (40 mL) before drying with anhydrous sodium sulphate and evaporation of solvents to yield a white powder. Purification was carried out by recrystalisation from chloroform to afford the target compound as colourless crystals (21.96 g, 52.01 mmol, 77 %).

^{1}H NMR (Bruker AV300 @ 300.22, $CDCl_3$, 298K) δ_H (ppm): 7.90 (d, *J* = 8.6 Hz, 2H, ArH), 7.63 (d, *J* = 8.6 Hz, 2H, ArH), 7.40 (d, *J* = 8.3 Hz, 2H, ArH), 7.30 (d, *J* = 8.3 Hz, 2H, ArH), 2.46 (s, 3H, ArMe)

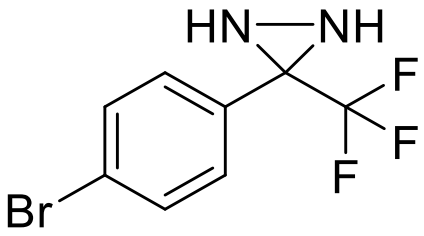
^{13}C NMR (Bruker AV300 @ 75.50, $CDCl_3$, 298K) δ_C (ppm): δ 152.90 (q, $^{2}J_{C\text{-}F}$ = 33.7 Hz), 146.32, 132.14, 130.79, 129.88, 129.87, 129.16, 126.48, 123.18, 119.31 (q, $^{1}J_{C\text{-}F}$ = 277.7 Hz), 21.67

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298K) δ_F (ppm): -66.73

HRMS (ESI-TOF) *m/z*: $[M+Na]^+$ Calculated for $C_{15}H_{11}BrF_3NNaO_3S$ 443.9493; Found 443.9493

^{1}H and ^{13}C NMR data match reported results.[85] A literature search yielded no previous report of ^{19}F NMR data.

3-(4-bromophenyl)-3-(trifluoromethyl)diaziridine (35)[85]



1-(4-Bromophenyl)-2,2,2-trifluoroethanone O-(*p*-tolylsulfonyl)oxime (5.00 g, 11.8 mmol, 1.0 equiv.) was dissolved in dichloromethane (100 mL) in a two-necked round bottom flask equipped with a dry ice condenser. The reaction vessel and condenser were cooled to -78°C. Gaseous ammonia was then introduced into the vessel and allowed to condense over a period of 60 minutes. The gas flow was then removed and the reaction mixture was allowed to stir for 18 hours. The condenser and reaction vessel were maintained at a temperature of -78°C for the first 6 hours before being allowed to warm to room temperature. Once all remaining ammonia had evaporated, the organic phase was washed with two portions of water and one of brine before drying with anhydrous magnesium sulphate. Evaporation of the solvent under reduced pressure gave a white solid (3.06 g, 11.5 mmol, 98 %).

^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 7.54 (d, *J* = 8.5 Hz, 2H, ArH), 7.47 (d, *J* = 8.5 Hz, 2H, ArH), 2.83 (d, *J* = 8.6 Hz, 1H, NH), 2.23 (d, *J* = 8.6 Hz, 1H, NH)

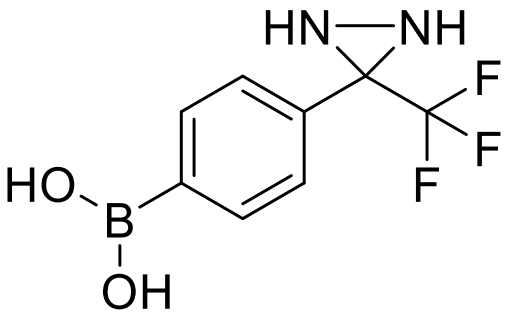
^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 132.01, 130.67, 129.75, 124.57, 123.28 (q, $^{1}J_{C\text{-}F}$ = 278.1 Hz), 57.62 (q, $^{2}J_{C\text{-}F}$ = 36.0)

^{19}F NMR (Agilent 500 @ 470.54 MHz, $CDCl_3$, 298 K) δ_F (ppm): -75.48

HRMS (ESI-TOF) *m/z*: $[M+H]^+$ Calculated for $C_8H_7BrF_3N_2$ 266.9745; Found 266.9746

^{1}H and ^{13}C NMR data match reported results.[135] A literature search yielded no previous report of ^{19}F NMR data.

4-(3-(trifluoromethyl)diaziridine-3-yl)phenylboronic acid (38)

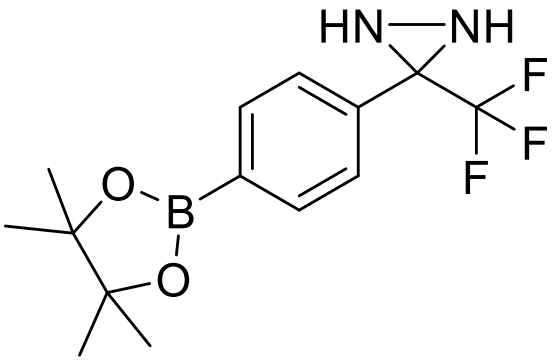


3-(4-Bromophenyl)-3-trifluoromethyl diaziridine (3.36g, 12.6 mmol, 1.0 equiv.) was dissolved in dichloromethane (60 mL) in a pre-dried schlenk flask. Triethylamine (10.5 mL, 75.3 mmol, 6.0 equiv.) was added and the mixture stirred for 1 hour before cooling to -78 °C. Once cooled trimethylsilyl trifluoromethanesulfonate (5.00 mL, 27.6 mmol, 2.2 equiv.) was added dropwise. Following complete addition the mixture was stirred for 1 hour at -78 °C, then warmed to room temperature and stirred for a further 3 hours at room temperature. The solvent was then evaporated under reduced pressure before the residue was extracted with two portions of anhydrous hexane (50 mL). The hexane solution was then cooled to -78 °C and a 2.0 M solution of n-butyl lithium in hexane (6.3 mL, 12.6 mmol, 1.0 equiv.) was added dropwise. Following the addition, the reaction was left to stir for 1 hour. Trimethyl boronate (1.4 mL, 12.6 mmol, 1.0 equiv.) was then added dropwise and the mixture was allowed to warm to room temperature before quenching with water. After separation, the organic layer was washed with two portions of water and one of brine. Evaporation of solvent gave an orange oil, the product was then used to form the pinacol ester without further purification.

^{1}H NMR (Bruker AV250 @ 250.13 MHz, $CDCl_3$, 298 K) δ_H (ppm): 2.29 (1H, d, *J* = 8.5 Hz, NH), 2.85 (1H, d, *J* = 8.5 Hz, NH), 7.72 (2H, d, *J* = 7.9 Hz, CH_{ar}) 8.21 (2H, d, *J* = 8.2 Hz, ArH)

^{11}B NMR (Bruker AV300 @ 96.32 MHz, $CDCl_3$, 298 K) δ_B (ppm): 34.71

3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine (39)



Method 1:

4-(3-(Trifluoromethyl)diaziridine-3-yl)phenylboronic acid (2.87 g, 12.4 mmol, 1.0 equiv.) and pinacol (1.61 g, 13.6 mmol, 1.0 equiv.) were suspended in chloroform and the mixture was allowed to stir for 2 hours. The mixture was then washed through a silica gel plug with diethylether. Evaporation of solvent under reduced pressure afforded the target compound as a white solid (2.82 g, 12.1 mmol, 97 %).

Method 2:

2,2,2-Trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxyborolan-2-yl)phenyl)ethan-1-one *O*-tosyl oxime (1.23 g, 2.62 mmol, 1.0 equiv.) was dissolved in dicholormethane (3 mL) in a two-necked round bottomed flask equipped with a dry ice condenser. The reaction vessel and condenser were cooled to -78 °C. Gaseous ammonia was then introduced into the vessel and allowed to condense over a period of 30 minutes. The gas flow was then removed and the reaction mixture was allowed to stir for 18 hours. The condenser and reaction vessel were maintained at a temperature of -78°C for the first 6 hours before being allowed to warm to room temperature. Once all remaining ammonia had evaporated, the organic phase was washed with two portions of water and one of brine before drying with anhydrous magnesium sulphate. Evaporation of the solvent under reduced pressure gave a white solid (0.661 g, 2.10 mmol, 80 %).

Method 3:

3-(4-Bromophenyl)-3-(trifluoromethyl)diaziridine (80.1 mg, 0.30 mmol, 1.0 equiv.), bis(pinacolato)diboron (91.4 mg, 0.36 mmol, 1.2 equiv.), potassium acetate (88.2 mg, 0.90 mmol, 3.0 equiv.), and [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium (II) dichloromethane complex (12.2 mg, 0,015 mmol, 0.05 equiv.) were added to a reaction vessel

and dissolved in dioxane (2.0 mL). The mixture was heated to 90 °C and stirred at this temperature for 18 hours. The mixture was cooled and filtered through silica with *tert*-butyl methyl ether, the filtrate was then washed with brine three portions of brine (10.0 mL). The organic layer was collected, then, following evaporation of the solvent, analysed by nuclear magnetic resonance spectroscopy (86 % conversion by NMR).

^{1}H NMR (Bruker AV300 @ 300.22 MHz, $CDCl_3$, 298 K) δ_H (ppm): 7.86 (d, *J* = 8.3 Hz, 2H, ArH), 7 .61 (d, *J* = 8.3 Hz, 2H, ArH), 2.84 (dd, *J* = 8.8, 1.1 Hz, 1H, NH), 2.31 (dd, *J* = 8.8, 1.1 Hz, 1H, NH), 1.34 (s, 12H, Me_2C)

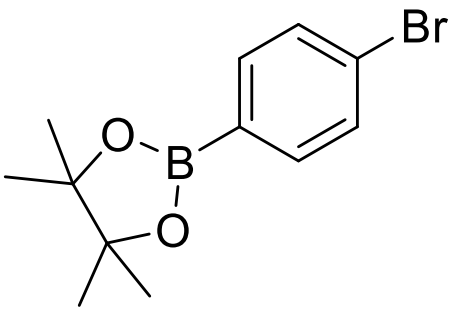
^{13}C NMR (Bruker AV300 @ 75.50 MHz, $CDCl_3$, 298 K) δ_C (ppm): 134.99, 134.20, 127.27, 123.42 (q, $^{1}J_{C\text{-}F}$ = 279.8 Hz), 84.04, 58.02 (q, $^{2}J_{C\text{-}F}$ = 36.0 Hz), 24.75

^{11}B NMR (Bruker AV300 @ 96.32 MHz, $CDCl_3$, 298 K) δ_B (ppm): 33.79

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298 K) δ_F (ppm): -75.42

HRMS (ESI-TOF) *m/z*: $[M+H]^+$ Calculated for $C_{14}H_{19}BF_3N_2O_2$ 315.1494; Found 315.1563

2-(4-bromophenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (40)[136]



Dibromobenzene (70.77 mg, 0.3 mmol, 1.0 equiv), bis(pinacolato)diboron (91.4 mg, 0.36 mmol, 1.2 equiv.), potassium acetate (88.2 mg, 0.90 mmol, 3.0 equiv.), and [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium (II) dichloromethane complex (12.2 mg, 0,015 mmol, 0.05 equiv.) were added to a reaction vessel and dissolved in dioxane (2.0 mL). The mixture was heated to 90 °C and stirred at this temperature for 18 hours. The mixture was cooled and filtered through silica with *tert*-butyl methyl ether, the filtrate was then washed with brine three portions of brine (10.0 mL). The organic layer was collected, then, following evaporation of the solvent, analysed by nuclear magnetic resonance spectroscopy (25 % conversion by NMR).

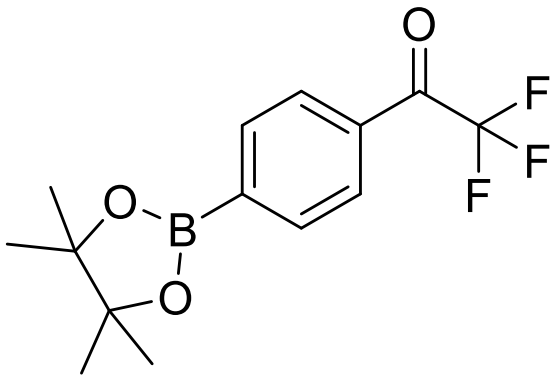
^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 7.62 (d, *J* = 8.3 Hz, 2H, ArH), 7.47 (d, *J* = 8.3 Hz, 2H, ArH), 1.31 (s, 12H, Me_2C)

^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 136.27, 133.83, 130.89, 67.01, 24.80

^{11}B NMR (Agilent 500 @ 160.44 MHz, $CDCl_3$, 298 K) δ_B (ppm): 22.24

These NMR data agree with the reported results.[136]

2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one (41)[137]



1-(4-Bromophenyl)-2,2,2-trifluoroethan-1-one (75.9 mg, 0.3 mmol, 1.0 equiv), bis(pinacolato)diboron (91.4 mg, 0.36 mmol, 1.2 equiv.), potassium acetate (88.2 mg, 0.90 mmol, 3.0 equiv.), and [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium (II) dichloromethane complex (12.2 mg, 0,015 mmol, 0.05 equiv.) were added to a reaction vessel and dissolved in dioxane (2.0 mL). The mixture was heated to 90 °C and stirred at this temperature for 18 hours. The mixture was cooled and filtered through silica with *tert*-butyl methyl ether, the filtrate was then washed with brine three portions of brine (10.0 mL). The organic layer was collected, then, following evaporation of the solvent, analysed by nuclear magnetic resonance spectroscopy (84 % conversion by NMR).

^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 8.03 (d, *J* = 8.4 Hz, 2H, ArH), 7.95 (d, *J* = 8.4 Hz, 2H, ArH), 1.34 (s, 12H, Me_2C)

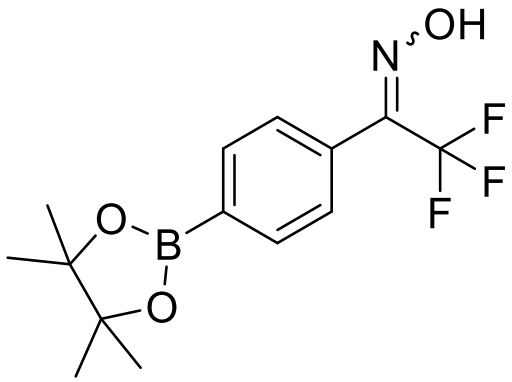
^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 180.79 (q, $^{2}J_{C-F}$ = 34.8 Hz), 135.16, 133.85, 128.89, 116.60 (q, $^{1}J_{C-F}$ = 290.8 Hz), 67.05, 24.82

^{11}B NMR (Agilent 500 @ 160.44 MHz, $CDCl_3$, 298 K) δ_B (ppm): 22.40

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298 K) δ_F (ppm): -71.50

^{1}H, ^{13}C, and ^{19}F NMR data matches the previously reported results.[138] A literature search yielded no previous report of the ^{11}B NMR data.

2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one oxime (42)



Method 1:

2,2,2-Trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one (0.150 g, 0.5 mmol, 1.0 equiv.) and hydroxylamine hydrochloride (34.7 mg, 0.5 mmol, 1.0 equiv.) were dissolved in ethanol (0.15 mL) and pyridine (0.30 mL), heated to 60 °C and stirred for 5 hours. After removal of the solvent under vacuum, the residue was dissolved in diethyl ether (5.0 mL) then washed with two portions of water (5.0 mL) and brine (5.0 mL). The organic phase was dried with anhydrous sodium sulphate. Evaporation of the solvent yield an orange oil. Purification by column chromatography afforded the target compound (0.137 g, 0.435 mmol, 87 %). R_f = 0.54 (hexane/EtOAc, 3:1).

Method 2:

1-(4-Bromophenyl)-2,2,2-trifluoroethan-1-one oxime (80.4 mg, 0.3 mmol, 1.0 equiv.), bis(pinacolato)diboron (91.4 mg, 0.36 mmol, 1.2 equiv.), potassium acetate (88.2 mg, 0.90 mmol, 3.0 equiv.), and [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium (II) dichloromethane complex (12.2 mg, 0,015 mmol, 0.05 equiv.) were added to a reaction vessel and dissolved in dioxane (2.0 mL), heated to 90 °C and stirred for 18 hours. After cooling, the mixture was filtered through a silica plug with *tert*-butyl methyl ether, the filtrate was washed with three portions of brine (10.0 mL). The organic layer was collected before analysis (24 % conversion by NMR).

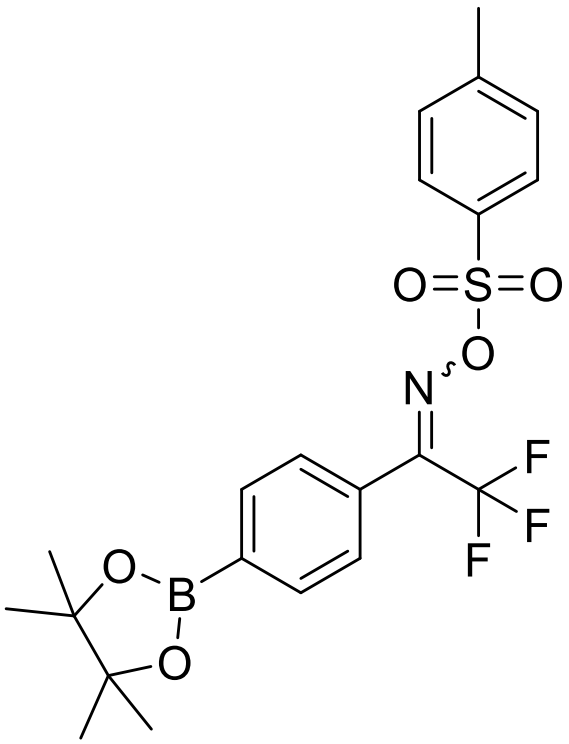
^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 10.90 (br s, 1H, NOH), 7.67 (d, *J* = 8.2 Hz, 2H, ArH), 7.47 (d, *J* = 8.2 Hz, 2H, ArH), 1.26 (s, 12H, Me_2C)

^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 146.80 (q, $^{2}J_{C-F}$ = 34.8 Hz), 134.59, 131.07, 127.47, 124.40 (q, $^{1}J_{C-F}$ = 282.6 Hz), 60.19, 26.80

^{11}B NMR (Agilent 500 @ 160.44 MHz, $CDCl_3$, 298 K) δ_B (ppm): 22.33

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298K) δ_F (ppm): -78.41

2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one *O*-tosyl oxime (43**)**



Method 1:

2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2yl)phenyl)ethan-1-one *O*-tosyl oxime (1.01 g, 3.20 mmol, 1.0 equiv.) was dissolved in dichloromethane (4.0 mL). Triethylamine (0.893 mL, 6.40 mmol, 2.0 equiv.) and 4-dimethylaminopyridine (0.039 g, 0.32 mmol, 0.1 equiv.) were added to the solution. The mixture was then cooled to 0 °C. After cooling, *p*-toluenesulfonyl chloride (0.723 g, 3.84 mmol, 1.2 equiv.) was added. The mixture was then allowed to warm to room temperature and then stirred for 3 hours. The solvent was removed under reduced pressure and the residue dissolved in diethyl ether (5.0 mL) and washed with two portions of water (5.0 mL) and one of brine (5.0 mL). The organic layer was dried with anhydrous sodium sulphate then the solvent evaporated under reduced pressure to yield a white powder. Recrystalisation from chloroform yielded the target compound as colourless crystals (1.23 g, 2.62 mmol, 82 %).

Method 2:

1-(4-bromophenyl)-2,2,2-trifluoroethan-1-one *O*-tosyl oxime (126 mg, 0.3 mmol, 1.0 equiv), bis(pinacolato)diboron (91.4 mg, 0.36 mmol, 1.2 equiv.), potassium acetate (88.2 mg, 0.90 mmol, 3.0 equiv.), and [1,1'-bis(diphenylphosphino)ferrocene]dichloropalladium (II) dichloromethane complex (12.2 mg, 0,015 mmol, 0.05 equiv.) were added to a reaction vessel

and dissolved in dioxane (2.0 mL). The mixture was heated to 90 °C and stirred at this temperature for 18 hours. The mixture was cooled and filtered through silica with *tert*-butyl methyl ether, the filtrate was then washed with brine three portions of brine (10.0 mL). The organic layer was collected, then, following evaporation of the solvent, analysed by nuclear magnetic resonance spectroscopy (54 % conversion by NMR).

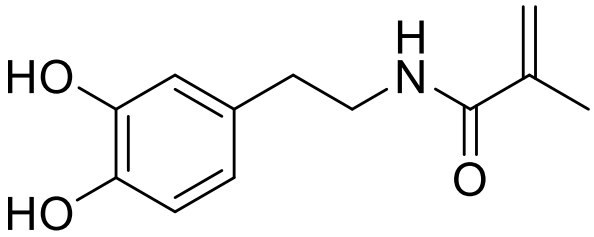
^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 7.94 (d, *J* = 8.5 Hz, 2H, ArH), 7.87 (d, *J* = 8.5 Hz, 2H, ArH), 7.33 (d, *J* = 8.3 Hz, 2H, ArH), 7.26 (d, *J* = 8.3 Hz, 2H, ArH), 2.05 (s, 3H, ArMe), 1.33 (s, 12H, Me_2C)

^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 146.19 (q, $^{2}J_{C\text{-}F}$ = 26.1 Hz), 135.15, 134.92, 132.23, 129.85, 129.21, 128.87, 127.33, 124.43 (q, $^{1}J_{C\text{-}F}$ = 288.9 Hz). 67.03, 24.77

^{11}B NMR (Agilent 500 @ 160.44 MHz, $CDCl_3$, 298 K) δ_B (ppm): 22.28

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298K) δ_F (ppm): -71.54

N-(3,4-dihydroxyphenethyl)methacrylamide (45)[119]



Sodium borate (2.41 g, 12.0 mmol, 0.8 equiv.) and sodium bicarbonate (1.89 g, 22.5 mmol, 1.5 equiv.) were dissolved in deionised water (50mL). The solution was then degassed using nitrogen for 45 minutes. Dopamine hydrochloride (2.84 g, 15.0 mmol, 1.0 equiv.) was added and the resulting mixture stirred under nitrogen. A solution of methyacrylic anhydride in degassed THF (1.27 M, 11.81 mL, 15.0 mmol, 1.0 equiv.) was added dropwise to the reaction mixture; the pH of the solution was kept between 8.0 and 9.0 by adding, as necessary, a 1.0 M sodium hydroxide solution in degassed deionised water. Once the addition was complete, the mixture was stirred under nitrogen at room temperature for 17 hours. The mixture was then washed with two portions of ethyl acetate (25 mL) and the aqueous layer filtered. A 6.0 M solution of HCl in deionised water was added to the filtrate until the pH reached 2.0, giving a light grey precipitate. The precipitate was extracted with three portions of ethyl acetate (25 mL), which were combined and condensed to approximately 2.0 mL. The solution was then precipitated into hexane at 0 °C and the precipitate collected by filtration. The product was further purified from boiling ethyl acetate to afford the desired compound as light grey crystals (2.00 g, 9.03 mmol, 60 %).

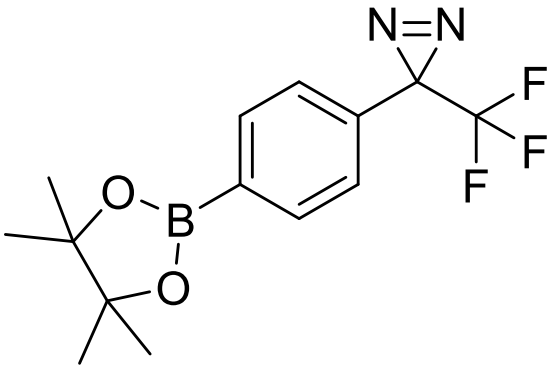
^{1}H NMR (Agilent 500 @ 500.06 MHz, $(CD_3)_2SO$, 298 K) δ_H (ppm): 7.94 (t, *J* = 5.5 Hz, 1H, NH), 7.91 (br, 2H, ArOH), 6.67 (d, *J* = 8.0 Hz, 1H, ArH (HOC-C**H**-CH)), 6.63 (d, *J* = 1.9 Hz, 1H, ArH (HOC-C**H**-C(R))), 6.46 (dd, *J* 8.0, 1.9 Hz, 1H, ArH (CH-C**H**-C(R))), 5.65 (s, 1H, R_2C=C**H$_2$**), 5.31 (t, *J* = 1.5 Hz, 1H, R_2C=C**H$_2$**), 3.30 (td, *J* = 7.7, 5.5 Hz, 2H, C**H$_2$**NH), 2.60 (t, *J* = 7.7 Hz, 2H, CH_2), 1.87 (s, 3H, CH_3)

^{13}C NMR (Agilent 500 @ 125.74 MHz, $(CD_3)_2SO$, 298 K) δ_C (ppm): 169.30, 146.89, 145.33, 141.91, 132.19, 121.12, 120.70, 117.86, 117.36, 42.84, 41.23, 36.44, 20.47

ESI-MS: *m/z* found 220.0971 $[M-H]^-$, expected value for $C_{12}H_{14}NO_3$ 220.0979

These NMR data match the reported results.[139]

3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine (46)



(4-(3-(trifluoromethyl)diaziridin-3-yl)phenyl)boronic acid pinacol ester (57 mg, 0.18 mmol, 1.0 equiv.) was dissolved in ethanol (0.5 mL), then trimethylamine (110.2 mg, 1.09 mmol, 6.0 equiv.) was added. A solution of iodine (20 mg/mL in ethanol, 0.079 M) was then added dropwise until the yellow colour persisted for 60 seconds. The reaction vessel was then left to stir in the dark for 30 mintues. The solvent was removed under reduced pressure, while ensuring the vessel was kept dark, the residue was then dissolved in diethyl ether and washed with 2 portions of an HCl solution in water (1.0 mM). The organic phase was dried with anhydrous sodium sulphate before evaporation of solvent gave the target compound as a white solid (52 mg, 0.17 mmol, 92 %).

^{1}H NMR (Agilent 500 @ 500.06 MHz, $CDCl_3$, 298 K) δ_H (ppm): 7.82 (d, *J* = 8.3 Hz, 2H, ArH), 7.17 (d, *J* = 8.3 Hz, 2H, ArH), 1.34 (s, 12H, CCH_3)

^{13}C NMR (Agilent 500 @ 125.74 MHz, $CDCl_3$, 298 K) δ_C (ppm): 135.02, 131.78, 125.49, 122.06 (q, $^{1}J_{C\text{-}F}$ = 269.0 Hz), 84.12, 28.5 (q, $^{2}J_{C\text{-}F}$ = 40.3 Hz), 24.80

^{19}F NMR (Agilent 500 @ 470.53 MHz, $CDCl_3$, 298K) δ_F(ppm): -65.10

7.3 Protein Labelling

Human Serum Albumin

To a solution of human serum albumin (1.0 mL, 0.030 µmol, 1.0 equiv., 2.0 mg/mL in water), a solution of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine was added (47.28 µL, 0.151 µmol, 5.0 equiv., 1.0 mg/mL in ethanol). The mixture was shaken for 60 seconds and then allowed to stand for 18 hours. Following this a solution of trimethylamine (91.39 µL, 30.0 equiv. 1.0 mg/mL in water) was added, then a solution of iodine (20 mg/mL in ethanol) was added dropwise until the colour persisted for 60 seconds. The mixture was then allowed to stand in the dark for 30 minutes, before exposure to UV light (354 nm) for 60 seconds. The sample was then analysed by mass spectrometry or gel electrophoresis without further purification.

Human Serum Albumin Non-Targeting Control

To a solution of human serum albumin (1.0 mL, 0.030 µmol, 1.0 equiv., 2.0 mg/mL in water), a solution of 3-(4-bromophenyl)-3-trifluoromethyl diaziridine was added (40.20 µL, 0.151 µmol, 5.0 equiv., 1.0 mg/mL in ethanol). The mixture was shaken for 60 seconds and then allowed to stand for 18 hours. Following this a solution of trimethylamine (91.39 µL, 30.0 equiv. 1.0 mg/mL in water) was added, then a solution of iodine (20 mg/mL in ethanol) was added dropwise until the colour persisted for 60 seconds. The mixture was then allowed to stand in the dark for 30 minutes, before exposure to UV light (354 nm) for 60 seconds. The sample was then analysed by mass spectrometry or gel electrophoresis without further purification.

Human Serum Albumin Non-Binding Control

To a solution of human serum albumin (1.0 mL, 0.030 µmol, 1.0 equiv., 2.0 mg/mL in water), a solution of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine was added (47.28 µL, 0.151 µmol, 5.0 equiv., 1.0 mg/mL in ethanol). The mixture was shaken for 60 seconds and then allowed to stand for 18 hours. Following this, the mixture was exposured to UV light (354 nm) for 60 seconds. The sample was then analysed by mass spectrometry or gel electrophoresis without further purification.

Casein

To a solution of casein (1.0 mL, 0.080 µmol, 1.0 equiv., 2.0 mg/mL in water), a solution of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine was added (125.6 µL, 0.40 µmol, 5.0 equiv., 1.0 mg/mL in ethanol). The mixture was shaken for 60 seconds and then allowed to stand for 18 hours. Following this a solution of trimethylamine

(242.9 µL, 30.0 equiv. 1.0 mg/mL in water) was added, then a solution of iodine (20 mg/mL in ethanol) was added dropwise until the colour persisted for 60 seconds. The mixture was then allowed to stand in the dark for 30 minutes, before exposure to UV light (354 nm) for 60 seconds. The sample was then analysed by mass spectrometry or gel electrophoresis without further purification.

Glycated Macrophage Migration Inhibitory Factor

To a solution of glycated macrophage migration inhibitory factor (0.4 mL, 0.027 µmol, 1.0 equiv., 0.87 mg/mL in water), a solution of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine was added (42.8 µL, 0.40 µmol, 5.0 equiv., 1.0 mg/mL in ethanol). The mixture was shaken for 60 seconds and then allowed to stand for 18 hours. Following this a solution of trimethylamine (82.8 µL, 30.0 equiv. 1.0 mg/mL in water) was added, then a solution of iodine (20 mg/mL in ethanol) was added dropwise until the colour persisted for 60 seconds. The mixture was then allowed to stand in the dark for 30 minutes, before exposure to UV light (354 nm) for 60 seconds. The sample was then analysed by mass spectrometry or gel electrophoresis without further purification.

Macrophage Migration Inhibitory Factor Non-Glycated Control

To a solution of macrophage migration inhibitory factor (0.4 mL, 0.075 µmol, 1.0 equiv., 2.4 mg/mL in water), a solution of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine was added (118.2 µL, 0.38 µmol, 5.0 equiv., 1.0 mg/mL in ethanol). The mixture was shaken for 60 seconds and then allowed to stand for 18 hours. Following this a solution of trimethylamine (228.4 µL, 30.0 equiv. 1.0 mg/mL in water) was added, then a solution of iodine (20 mg/mL in ethanol) was added dropwise until the colour persisted for 60 seconds. The mixture was then allowed to stand in the dark for 30 minutes, before exposure to UV light (354 nm) for 60 seconds. The sample was then analysed by mass spectrometry or gel electrophoresis without further purification.

7.4 Electrochemistry

Glassy Carbon Electrode

Experiment 1

A solution of potassium nitrate in water (0.1 M, 10.0 mL) was placed inside the electrochemical cell. The glassy carbon electrode (3 mm diameter) and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate to give a baseline reading.

3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine (3.1 mg, 9.9 µmol) was then dissolved into the electrolyte and further cyclic voltammograms obtained at room temperature with a 100 mVs^{-1} scan rate.

Experiment 2

A phosphate buffer solution (pH 7.2, 10.0 mL) was placed inside the electrochemical cell. The glassy carbon electrode (3 mm diameter) and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate to give a baseline reading.

The electrochemical cell was then shielded from sunlight by wrapping the cell in aluminium foil before 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine (1.7 mg, 9.9 µmol) was then dissolved into the electrolyte and further cyclic voltammograms obtained at room temperature with a 100 mVs^{-1} scan rate. The aluminium foil was then removed and the cell exposed to sunlight before obtaining another set of cyclic voltammograms with the same parameters.

Experiment 3

A phosphate buffer solution (pH 10.4, 10.0 mL) was placed inside the electrochemical cell. The glassy carbon electrode (3 mm diameter) and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate to give a baseline reading.

3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl)diaziridine (1.7 mg, 9.9 µmol) was then dissolved into the electrolyte and further cyclic voltammograms obtained at room temperature with a 100 mVs^{-1} scan rate.

Experiment 4

A solution of tetrabutylammonium hexafluorophoshate in acetonitrile (0.1 M, 10.0 mL) was placed inside the electrochemical cell. The glassy carbon electrode (3 mm diameter) and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate to give a baseline reading.

3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine (3.0 mg, 9.6 µmol) was then dissolved into the electrolyte and further cyclic voltammograms obtained at room temperature with a 100 mVs^{-1} scan rate. A further portion of 3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine (3.0 mg, 9.6 µmol) was added to give a total of 6.0 mg (19.2 µmol) and cyclic voltammograms were then obtained at a range of scan rates (100 mVs^{-1}, 200 mVs^{-1}, 500 mVs^{-1}, 1000 mVs^{-1}, 2000 mVs^{-1})

Fluorine Doped Tin Oxide

Experiment 1

3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine (6.0 mg, 19.2 µmol) was dissolved a solution of tetrabutylammonium hexafluorophoshate in acetonitrile (0.1 M, 10.0 mL) in the electrochemical cell, the cell was then wrapped in aluminium foil to shield from the light. The untreated fluorine doped tin oxide electrode and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate. Once completed, the entire cell was uncovered and placed underneath an ultraviolet lamp used for visualisation of thin layer chromatography plates. The fluorine doped tin oxide electrode was then removed from the cell and set aside to be used at the treated electrode.

Experiment 2

A solution of Alizarin Red S in phosphate buffer (0.1 M, pH 7.2, 10.0 mL) was placed inside the electrochemical cell. The untreated fluorine doped tin oxide electrode and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate.

Experiment 3

A solution of Alizarin Red S in phosphate buffer (0.1 M, pH 7.2, 10.0 mL) was placed inside the electrochemical cell. The fluorine doped tin oxide electrode treated with 3-(4-(4,4,5,5-

tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine and platinum wire counter electrode were then inserted into the solution. Cyclic voltammograms were then performed at room temperature with a 100 mVs^{-1} scan rate.

Chapter 8: References

1. R. C. Hughes, *Glycoproteins*, London : Chapman and Hall, London, 1983.
2. M. Nagae and Y. Yamaguchi, *Int. J. Mol. Sci.*, 2012, **13**, 8398-8429.
3. A. Varki, R. Cummings, J. Esko, H. Freeze, G. Hart and J. Marth, *Essentials of glycobiology*, Cold Spring Harbor, N.Y. : Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y., 2009.
4. J. Hofsteenge, D. R. Muller, T. Debeer, A. Loffler, W. J. Richter and J. F. G. Vliegenthart, *Biochemistry*, 1994, **33**, 13524-13530.
5. M. E. Taylor, *Introduction to glycobiology*, Oxford : Oxford University Press, Oxford, 2006.
6. F. Schwarz and M. Aebi, *Curr. Opin. Struct. Biol.*, 2011, **21**, 576-582.
7. P. Van den Steen, P. M. Rudd, R. A. Dwek and G. Opdenakker, *Crit. Rev. Biochem. Mol. Biol.*, 1998, **33**, 151-208.
8. R. K. Murray, *Harper's Illustrated Biochemistry*, New York : McGraw-Hill Medical, New York, 28 edn., 2009.
9. T. Katoh and M. Tiemeyer, *Glycoconjugate J.*, 2013, **30**, 57-66.
10. R. Shogren, T. A. Gerken and N. Jentoft, *Biochemistry*, 1989, **28**, 5525-5536.
11. P. M. Rudd, T. Elliott, P. Cresswell, I. A. Wilson and R. A. Dwek, *Science*, 2001, **291**, 2370-2376.
12. B. Imperiali and S. E. O'Connor, *Curr. Opin. Chem. Biol.*, 1999, **3**, 643-649.
13. P. M. Rudd, M. R. Wormald, R. L. Stanfield, M. D. Huang, N. Mattsson, J. A. Speir, J. A. DiGennaro, J. S. Fetrow, R. A. Dwek and I. A. Wilson, *J. Mol. Biol.*, 1999, **293**, 351-366.
14. D. R. Madden, J. C. Gorga, J. L. Strominger and D. C. Wiley, *Cell*, 1992, **70**, 1035-1048.
15. R. Singh, A. Barden, T. Mori and L. Beilin, *Diabetologia*, 2001, **44**, 129-146.
16. S. Seo, S. Karboune, L. L'Hocine and V. Yaylayan, *Food Sci. Technol. LWB*, 2013, **53**, 44-53.
17. J. D. McPherson, B. H. Shilton and D. J. Walton, *Biochemistry*, 1988, **27**, 1901-1907.
18. E. Brinkmann, K. J. Wellsknecht, S. R. Thorpe and J. W. Baynes, *J Chem. Soc., Perkin Trans. 1*, 1995, 2817-2818.
19. S. J. Meade, A. G. Miller and J. A. Gerrard, *Bioorg. Med. Chem.*, 2003, **11**, 853-862.
20. R. Pamplona, *Chem.-Biol. Interact.*, 2011, **192**, 14-20.
21. G. Vistoli, D. De Maddis, A. Cipak, N. Zarkovic, M. Carini and G. Aldini, *Free Radical Res.*, 2013, **47 Suppl 1**, 3-27.
22. H. Vicente Miranda and T. F. Outeiro, *J. Pathol.*, 2010, **221**, 13-25.
23. M. Zoltowska, E. Delvin, E. Ziv, N. Peretti, M. Chartre and E. Levy, *Lipids*, 2004, **39**, 81-85.
24. G. Münch, B. Westcott, T. Menini and A. Gugliucci, *Amino Acids*, 2012, **42**, 1221-1236.
25. I. Gonzalez, J. Romero, B. L. Rodriguez, R. Perez-Castro and A. Rojas, *Immunobiology*, 2013, **218**, 790-797.
26. A. Bierhaus, P. M. Humpert, M. Morcos, T. Wendt, T. Chavakis, B. Arnold, D. M. Stern and P. P. Nawroth, *J. Mol. Med.*, 2005, **83**, 876-886.
27. R. Milne and S. Brownstein, *Amino Acids*, 2013, **44**, 1397-1407.
28. A. Rojas and M. A. Morales, *Life Sci.*, 2004, **76**, 715-730.
29. N. Ahmed, *Diabetes Research and Clinical Practice*, 2005, **67**, 3-21.
30. K. J. WellsKnecht, E. Brinkmann, M. C. WellsKnecht, J. E. Litchfield, M. U. Ahmed, S. Reddy, D. V. Zyzak, S. R. Thorpe and J. W. Baynes, *Nephrol. Dia. Transpl.*, 1996, **11**, 41-47.
31. S. Vasan, X. Zhang, X. N. Zhang, A. Kapurniotu, J. Bernhagen, S. Teichberg, J. Basgen, D. Wagle, D. Shih, I. Terlecky, R. Bucala, A. Cerami, J. Egan and P. Ulrich, *Nature*, 1996, **382**, 275-278.

32. M. Neeper, A. M. Schmidt, J. Brett, S. D. Yan, F. Wang, Y. C. E. Pan, K. Elliston, D. Stern and A. Shaw, *J. Biol. Chem.*, 1992, **267**, 14998-15004.
33. A. Bierhaus, S. Schiekofer, M. Schwaninger, M. Andrassy, P. M. Humpert, J. Chen, M. Hong, T. Luther, T. Henle, I. Kloting, M. Morcos, M. Hofmann, H. Tritschler, B. Weigle, M. Kasper, M. Smith, G. Perry, A. M. Schmidt, D. M. Stern, H. U. Haring, E. Schleicher and P. P. Nawroth, *Diabetes*, 2001, **50**, 2792-2808.
34. J. Amir, M. Waite, J. Tobler, D. L. Catalfamo, T. Koutouzis, J. Katz and S. M. Wallet, *Cell. Immunol.*, 2011, **272**, 45-52.
35. R. Bucala, K. J. Tracey and A. Cerami, *J. Clin. Invest.*, 1991, **87**, 432-438.
36. S. Yamagishi, H. Fujimori, H. Yonekura, Y. Yamamoto and H. Yamamoto, *Diabetologia*, 1998, **41**, 1435-1441.
37. H. J. Luth, V. Ogunlade, B. Kuhla, R. Kientsch-Engel, P. Stahl, J. Webster, T. Arendt and G. Munch, *Cereb. Cortex*, 2005, **15**, 211-220.
38. A. Andersson, R. Covacu, D. Sunnemark, A. I. Danilov, A. Dal Bianco, M. Khademi, E. Wallstrom, A. Lobell, L. Brundin, H. Lassmann and R. A. Harris, *J. Leukocyte Biol.*, 2008, **84**, 1248-1255.
39. G. Munch, H. J. Luth, A. Wong, T. Arendt, E. Hirsch, R. Ravid and P. Riederer, *J. Chem. Neuroanat.*, 2000, **20**, 253-257.
40. N. Sasaki, M. Takeuchi, H. Chowei, S. Kikuchi, Y. Hayashi, N. Nakano, H. Ikeda, S. Yamagishi, T. Kitamoto, T. Saito and Z. Makita, *Neurosci. Lett.*, 2002, **326**, 117-120.
41. L. Engelen, C. D. A. Stehouwer and C. G. Schalkwijk, *Diabetes Obes. Metab.*, 2013, **15**, 677-689.
42. S.-i. Yamagishi, *Rejuvenation Res.*, 2012, **15**, 564-572.
43. R. Inagi, *Methods Enzymol.*, 2011, **491**, 361-380.
44. S.-W. Lee, K.-H. Park, S. Park, J.-H. Kim, S.-Y. Hong, S.-K. Lee, D. Choi and Y.-B. Park, *Arthritis Rheum.*, 2013, **65**, 1902-1912.
45. S. Hollenbach, P. Thampi, T. Viswanathan and E. C. Abraham, *Mol. Cell. Biochem.*, 2003, **243**, 73-80.
46. T. Krechler, M. Jachymova, O. Mestek, A. Zak, T. Zima and M. Kalousova, *Clin. Biochem.*, 2010, **43**, 882-886.
47. T. Soulis, M. E. Cooper, D. Vranes, R. Bucala and G. Jerums, *Kidney Int.*, 1996, **50**, 627-634.
48. W. K. Bolton, D. C. Cattran, M. E. Williams, S. G. Adler, G. B. Appel, K. Cartwright, P. G. Foiles, B. I. Freedman, P. Raskin, R. E. Ratner, B. S. Spinowitz, F. C. Whittier, J. P. Wuerth and A. I. I. Grp, *Am. J. Nephrol.*, 2004, **24**, 32-40.
49. B. I. Freedman, J. P. Wuerth, K. Cartwright, R. P. Bain, S. Dippe, K. Hershon, A. D. Mooradian, B. S. Spinowitz and A. I. S. Investigators, *Controlled Clin. Trials*, 1999, **20**, 493-510.
50. D. A. Kass, E. P. Shapiro, M. Kawaguchi, A. R. Capriotti, A. Scuteri, R. C. deGroof and E. G. Lakatta, *Circulation*, 2001, **104**, 1464-1470.
51. J. W. L. Hartog, S. Willemsen, D. J. van Veldhuisen, J. L. Posma, L. M. van Wijk, Y. M. Hummel, H. L. Hillege, A. A. Voors and B. Investigators, *Eur. J. Heart Fail.*, 2011, **13**, 899-908.
52. W. C. Little, M. R. Zile, D. W. Kitzman, W. G. Hundley, T. X. O'Brien and R. C. Degroof, *J. Card. Fail.*, 2005, **11**, 191-195.
53. Y. Dong, M. Zhang, S. Wang, B. Liang, Z. Zhao, C. Liu, M. Wu, H. C. Choi, T. J. Lyons and M.-H. Zou, *Diabetes*, 2010, **59**, 1386-1396.
54. Y. Izuhara, M. Nangaku, S. Takizawa, S. Takahashi, J. Shao, H. Oishi, H. Kobayashi, C. v. Y. de Strihou and T. Miyata, *Nephrol. Dia. Transpl.*, 2008, **23**, 497-509.
55. M. E. Williams, W. K. Bolton, R. G. Khalifah, T. P. Degenhardt, R. J. Schotzinger and J. B. McGill, *Am. J. Nephrol.*, 2007, **27**, 605-614.

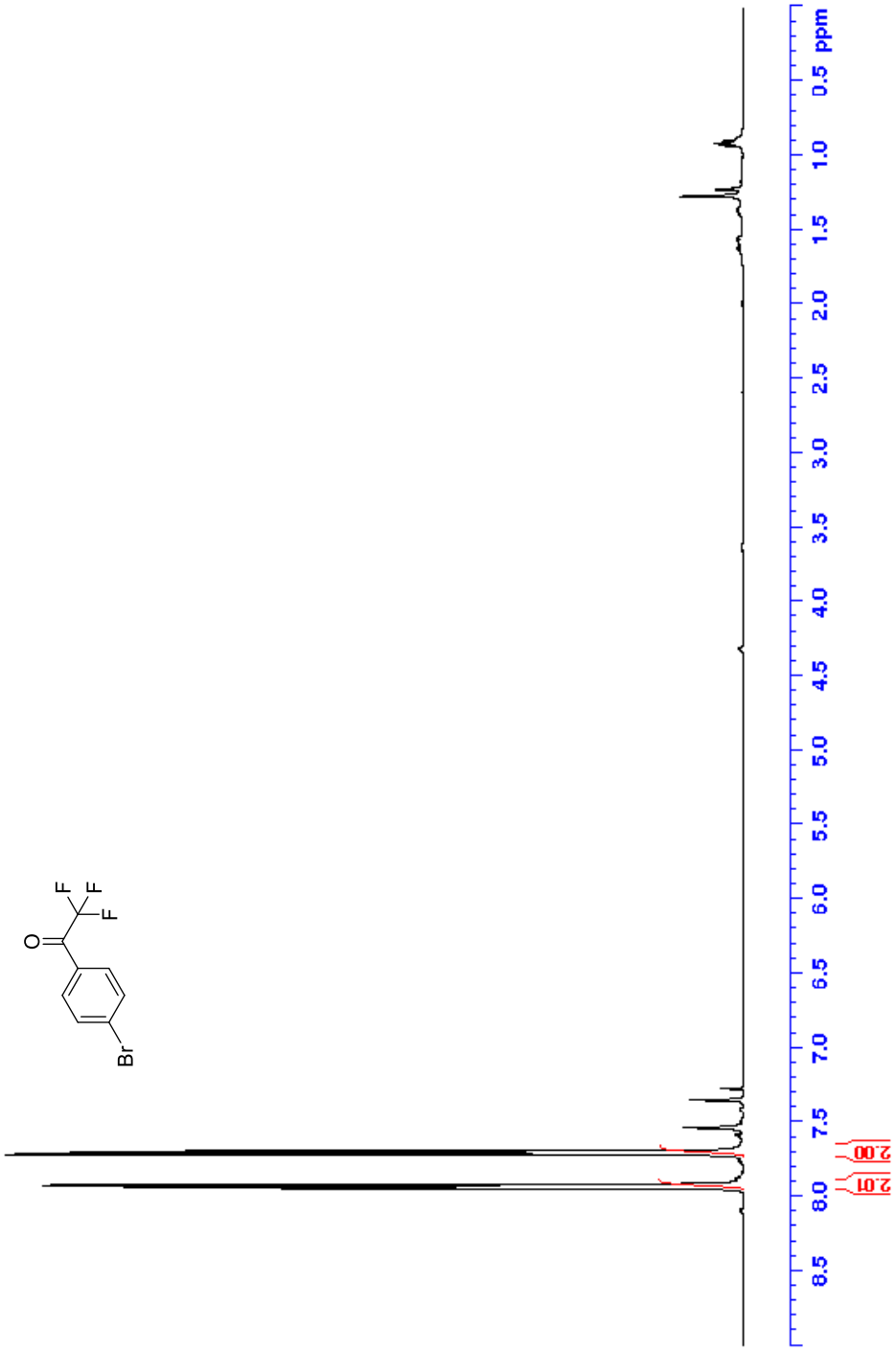
56. B. I. Hudson, L. G. Bucciarelli, T. Wendt, T. Sakaguchi, E. Lalla, W. Qu, Y. Lu, L. Lee, D. M. Stern, Y. Naka, R. Ramasamy, S. D. Yan, S. F. Yan, V. D'Agati and A. M. Schmidt, *Arch. Biochem. Biophys.*, 2003, **419**, 80-88.
57. T. Sakaguchi, S. F. Yan, S. Du Yan, D. Belov, L. L. Rong, M. Sousa, M. Andrassy, S. P. Marso, S. Duda, B. Arnold, B. Liliensiek, P. P. Nawroth, D. M. Stern, A. M. Schmidt and Y. Naka, *J. Clin. Invest.*, 2003, **111**, 959-972.
58. L. Park, K. G. Raman, K. J. Lee, Y. Lu, L. J. Ferran, W. S. Chow, D. Stern and A. M. Schmidt, *Nat. Med.*, 1998, **4**, 1025-1031.
59. T. Kislinger, N. Tanji, T. Wendt, W. Qu, Y. Lu, L. J. Ferran, A. Taguchi, K. Olson, L. Bucciarelli, M. Goova, M. A. Hofmann, G. Cataldegirmen, V. D'Agati, M. Pischetsrieder, D. M. Stern and A. M. Schmidt, *Arterioscler. Thromb. Vasc. Biol.*, 2001, **21**, 905-910.
60. A. Bierhaus, K. M. Haslbeck, P. M. Humpert, B. Liliensiek, T. Dehmer, M. Morcos, A. A. R. Sayed, M. Andrassy, S. Schiekofer, J. G. Schneider, J. B. Schulz, D. Heuss, B. Neundorfer, S. Dierl, J. Huber, H. Tritschler, A. M. Schmidt, M. Schwaninger, H. U. Haering, E. Schleicher, M. Kasper, D. M. Stern, B. Arnold and P. P. Nawroth, *J. Clin. Invest.*, 2004, **114**, 1741-1751.
61. M. T. Goova, J. Li, T. Kislinger, W. Qu, Y. Lu, L. G. Bucciarelli, S. Nowygrod, B. M. Wolf, X. Caliste, S. F. Yan, D. M. Stern and A. M. Schmidt, *Am. J. Pathol.*, 2001, **159**, 513-525.
62. J. M. Lehn, *Angew. Chem. Int. Ed.*, 1988, **27**, 89-112.
63. A. Hulanicki, S. Glab and F. Ingman, *Pure Appl. Chem.*, 1991, **63**, 1247-1250.
64. T. D. James and S. Shinkai, in *Host-Guest Chemistry: Mimetic Approaches to Study Carbohydrate Recognition*, ed. S. Penades, 2002, vol. 218, pp. 159-200.
65. T. D. James, M. D. Phillips and S. Shinkai, *Boronic Acids in Saccharide Recognition*, Royal Society of Chemistry, 2006.
66. M. D. Phillips, T. M. Fyles, N. P. Barwell and T. D. James, *Chem. Commun.*, 2009, 6557-6559.
67. T. D. James, K. Sandanayake, R. Iguchi and S. Shinkai, *J. Am. Chem. Soc.*, 1995, **117**, 8982-8987.
68. S. Jin, C. Zhu, Y. Cheng, M. Li and B. Wang, *Bioorg. Med. Chem.*, 2010, **18**, 1449-1455.
69. J. Arnaud, A. Audfray and A. Imberty, *Chem. Soc. Rev.*, 2013, **42**, 4798-4813.
70. R. Meerwaldt, R. Graaff, P. H. N. Oomen, T. P. Links, J. J. Jager, N. L. Alderson, S. R. Thorpe, J. W. Baynes, R. O. B. Gans and A. J. Smit, *Diabetologia*, 2004, **47**, 1324-1330.
71. D. S. Hage, J. A. Anguizola, C. Bi, R. Li, R. Matsuda, E. Papastavros, E. Pfaunmiller, J. Vargas and X. Zheng, *J. Pharm. Biomed. Anal.*, 2012, **69**, 93-105.
72. T. R. Jackson, J. S. Springall, D. Rogalle, N. Masurnoto, H. C. Li, F. D'Hooge, S. P. Perera, A. T. A. Jenkins, T. D. James, J. S. Fossey and J. M. H. van den Elsen, *Electrophoresis*, 2008, **29**, 4185-4191.
73. M. P. P. Morais, D. Marshall, S. E. Flower, C. J. Caunt, T. D. James, R. J. Williams, N. R. Waterfield and J. M. H. van den Elsen, *Sci. Rep.*, 2013, **3**.
74. D. J. Lapinsky, *Bioorg. Med. Chem.*, 2012, **20**, 6237-6247.
75. S. A. Sieber, P. Geurink, L. Prely, G. Marel, R. Bischoff and H. Overkleeft, in *Activity-Based Protein Profiling*, Springer Berlin Heidelberg, 2012, vol. 324, pp. 85-113.
76. A. M. Sadaghiani, S. H. L. Verhelst and M. Bogyo, *Curr. Opin. Chem. Biol.*, 2007, **11**, 20-28.
77. S.-H. Yu, A. M. Wands and J. J. Kohler, *J. Carbohydr. Chem.*, 2012, **31**, 325-352.
78. M. van Scherpenzeel, R. J. B. H. N. van den Berg, W. E. Donker-Koopman, R. M. J. Liskamp, J. M. F. G. Aerts, H. S. Overkleeft and R. J. Pieters, *Bioorg. Med. Chem.*, 2010, **18**, 267-273.
79. M. van Scherpenzee, E. E. Moret, L. Ballell, R. M. J. Liskamp, U. J. Nilsson, H. Leffler and R. J. Pieters, *ChemBioChem*, 2009, **10**, 1724-1733.
80. M. N. Gandy, A. W. Debowski and K. A. Stubbs, *Chem. Commun.*, 2011, **47**, 5037-5039.

81. A. V. Romaniouk, A. Silva, J. Feng and I. K. Vijay, *Glycobiology*, 2004, **14**, 301-310.
82. J. Z. Zhao and T. D. James, *J. Mater. Chem.*, 2005, **15**, 2896-2901.
83. J. Z. Zhao, T. M. Fyles and T. D. James, *Angew.Chem. Int. Ed.*, 2004, **43**, 3461-3464.
84. J. Z. Zhao, M. G. Davidson, M. F. Mahon, G. Kociok-Kohn and T. D. James, *J. Am. Chem. Soc.*, 2004, **126**, 16179-16186.
85. T. Bender, M. Huss, H. Wieezorek, S. Grond and P. von Zezschwitz, *Eur. J. Org. Chem.*, 2007, 3870-3878.
86. G. Deng, T. D. James and S. Shinkai, *J. Am. Chem. Soc.*, 1994, **116**, 4567-4572.
87. C. S. Huang, L. P. Wen, H. B. Liu, Y. L. Li, X. F. Liu, M. J. Yuan, J. Zhai, L. Jiang and D. B. Zhu, *Adv. Mater.*, 2009, **21**, 1721-1725.
88. J. P. Lorand and J. O. Edwards, *J. Org. Chem.*, 1959, **24**, 769-774.
89. T. Ishiyama, M. Murata and N. Miyaura, *J. Org. Chem.*, 1995, **60**, 7508-7510.
90. L. H. Wang, J. Y. Li, X. L. Cui, Y. S. Wu, Z. W. Zhu and Y. J. Wu, *Adv. Synth. Catal.*, 2010, **352**, 2002-2010.
91. G. Li, K. Y. Yoon, X. J. Zhong, X. Y. Zhu and G. B. Dong, *Chem.-Eur. J.*, 2016, **22**, 9116-9120.
92. S. Pospiech, M. Bolte, H. W. Lerner and M. Wagner, *Chem.-Eur. J.*, 2015, **21**, 8229-8236.
93. J. Frahn and A. D. Schluter, *Synthesis*, 1997, 1301-1304.
94. S. Jung, Y. Kitajima, Y. Ueda, K. Suzuki and K. Ohmori, *Synlett*, 2016, **27**, 1521-1526.
95. K. M. Dawood, M. R. Shaaban, M. B. Elamin and A. M. Farag, *Arab. J. Chem.*, 2017, **10**, 473-479.
96. J. Song, H. Zhao, Y. Liu, H. Han, Z. Li, W. Chu and Z. Sun, *New J. Chem.*, 2017, **41**, 372-376.
97. J. E. Baldwin, A. J. Pratt and M. G. Moloney, *Tetrahedron*, 1987, **43**, 2565-2575.
98. Y. Wang, H. Y. Yu, X. L. Shi, Z. P. Luo, D. H. Lin and M. D. Huang, *J. Biol. Chem.*, 2013, **288**, 15980-15987.
99. J. Anguizola, R. Matsuda, O. S. Barnaby, K. S. Hoy, C. L. Wa, E. DeBolt, M. Koke and D. S. Hage, *Clin. Chim. Acta*, 2013, **425**, 64-76.
100. A. Raghav and J. Ahmad, *Diabetes Metab. Syndr.*, 2014, **8**, 245-251.
101. Y. Liao, R. Alvarado, B. Phinney and B. Loennerdal, *J. Proteome Res.*, 2011, **10**, 5409-5415.
102. M. Inagaki, H. Muranishi, K. Yamada, K. Kakehi, K. Uchida, T. Suzuki, T. Yabe, T. Nakagomi, O. Nakagomi and Y. Kanamaru, *J.Dairy Sci.*, 2014, **97**, 2653-2661.
103. S. Jindal and A. Naeem, *J. Fluoresc.*, 2013, **23**, 367-374.
104. G. Hu, Y. Zheng, Z. Liu, Y. Xiao, Y. Deng and Y. Zhao, *Food Chem.*, 2017, **221**, 1860-1866.
105. G. Hu, Y. Zheng, Z. Liu, Y. Deng and Y. Zhao, *Food Chem.*, 2016, **204**, 46-55.
106. D. Scheidegger, G. Larsen and S. Clara Kivatinitz, *Int. Dairy J.*, 2016, **55**, 64-71.
107. D. Scheidegger, R. P. Pecora, P. M. Radici and S. C. Kivatinitz, *J. Dairy Sci.*, 2010, **93**, 5101-5109.
108. H. W. Sun, J. Bernhagen, R. Bucala and E. Lolis, *Proc. Natl. Acad. Sci. U. S. A.*, 1996, **93**, 5191-5196.
109. J. A. Chesney and R. A. Mitchell, *Mol. Med.*, 2015, **21**, S19-S24.
110. V. Richard, N. Kindt and S. Saussez, *Int. J. Oncol.y*, 2015, **47**, 1627-1633.
111. M. C. Morrison and R. Kleemann, *Front. Immunol.*, 2015, **6**, 1-13.
112. Y. I. Sanchez-Zamora and M. Rodriguez-Sosa, *J. Diabetes Res.*, 2014, **2014**, 804519.
113. A. Bruchfeld, M. Wendt and E. J. Miller, *Front. Immunol.*, 2016, **7**, 1-7.
114. J. Bloom and Y. Al-Abed, *J. Neuroinflammation*, 2014, **11**, 1-9.
115. M. F. Leyton-Jaimes, J. Kahn and A. Israelson, *Exp. Neurol.*, 2017.
116. O. Kassaar, M. P. Morais, S. Xu, E. L. Adam, R. C. Chamberlain, B. Jenkins, T. James, P. T. Francis, S. Ward, R. J. Williams and J. van den Elsen, *Sci. Rep.*, 2017, **7**, 42874.

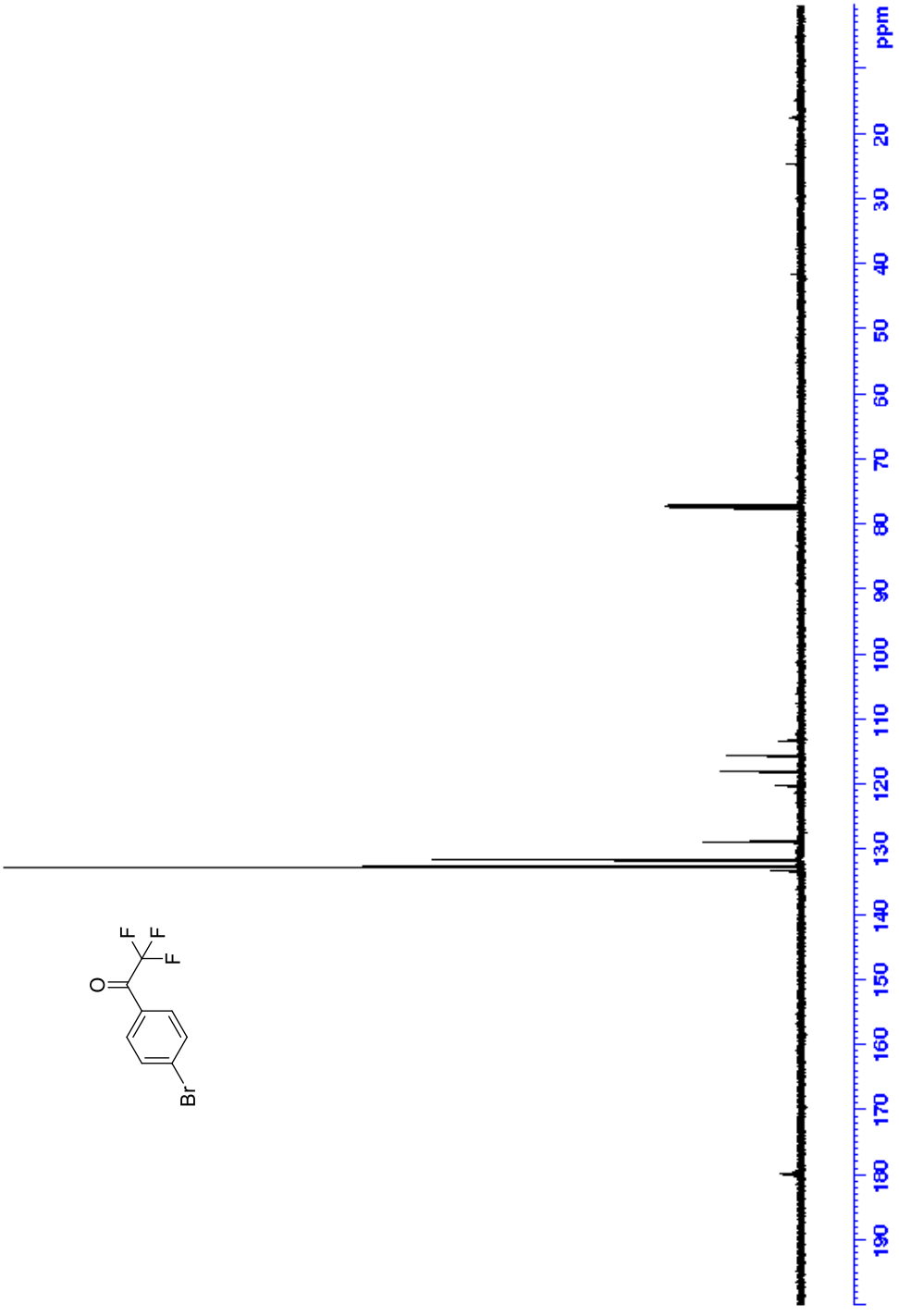
117. G. Grieb, B.-S. Kim, D. Simons, J. Bernhagen and N. Pallua, *Mini-Rev. Med. Chem.*, 2014, **14**, 1125-1131.
118. J. D. Tipton, J. C. Tran, A. D. Catherman, D. R. Ahlf, K. R. Durbin and N. L. Kelleher, *J. Biol. Chem.*, 2011, **286**, 25451-25458.
119. E. M. White, J. E. Seppala, P. M. Rushworth, S. Sharma and J. Locklin, *Abstr. Pap. Am. Chem. S.*, 2014, **247**.
120. K. Lawrence, S. E. Flower, G. Kociok-Kohn, C. G. Frost and T. D. James, *Anal. Methods*, 2012, **4**, 2215-2217.
121. G. Springsteen and B. H. Wang, *Chem. Commun.*, 2001, 1608-1609.
122. M. Vukicevic and H. H. Tonnesen, *Pharm. Dev. Technol.*, 2016, **21**, 428-436.
123. U. Guth, W. Vonau and J. Zosel, *Meas. Sci. Technol.*, 2009, **20**.
124. U. Akiba and J.-i. Anzai, *Sensors*, 2016, **16**, 2045/1-2045/18.
125. S. A. Zaidi and J. H. Shin, *Talanta*, 2016, **149**, 30-42.
126. A. Matsumoto, N. Sato, K. Kataoka and Y. Miyahara, *J. Am. Chem. Soc.*, 2009, **131**, 12022-12023.
127. Q. Wang, I. Kaminska, J. Niedziolka-Jonsson, M. Opallo, M. Li, R. Boukherroub and S. Szunerits, *Biosens. Bioelectron.*, 2013, **50**, 331-337.
128. C. M. Elson and M. T. H. Liu, *J. Chem. Soc., Chem. Commun.*, 1982, 415-416.
129. F. Lamberti, S. Agnoli, L. Brigo, G. Granozzi, M. Giomo and N. Elvassore, *ACS Appl. Mater. Inter.*, 2013, **5**, 12887-12894.
130. S. Shariki, O. T. L. Cox, D. A. Tickell, M. P. P. Morais, J. M. H. van den Elsen, T. D. James, S. E. C. Dale, S. Bending and F. Marken, *J. Mater. Chem.*, 2012, **22**, 18999-19006.
131. W. L. Zhang, X. L. Niu, H. Zhang, Q. Q. Jiang and P. P. Zhang, *Int, J. Electrochem.*, 2015, 1-6.
132. N. Tanjila, A. Rayhan, M. S. Alam, I. A. Siddiquey and M. A. Hasnat, *RSC Adv.*, 2016, **6**, 93162-93168.
133. M. D. Phillips and T. D. James, *J. Fluoresc.*, 2004, **14**, 549-559.
134. B. Cui, H. Sun, Y. B. Xu, L. L. Duan and Y. M. Li, *Tetrahedron*, 2017, **73**, 6754-6762.
135. I. Protasova, B. Bulat, N. Jung and S. Brase, *Org. Lett.*, 2017, **19**, 34-37.
136. L. Zhang and L. Jiao, *J. Am. Chem. Soc.*, 2017, **139**, 607-610.
137. D. Chen, C. L. Franklin, P. R. Guzzo, u. S. Lin, J. Liu, M. M. C. Lo, R. P. Nargund and I. K. Sebhat, *Official Gazette of the United States Patent and Trademark Office Patents*, 2012.
138. Y. Kadoh, M. Tashiro, K. Oisaki and M. Kanai, *Adv. Synth. Catal.*, 2015, **357**, 2193-2198.
139. H. O. Ham, Z. Liu, K. H. A. Lau, H. Lee and P. B. Messersmith, *Angew. Chem. Int. Edit.*, 2011, **50**, 732-736.

Appendix I: Nuclear Magnetic Resonance Spectroscopy Data

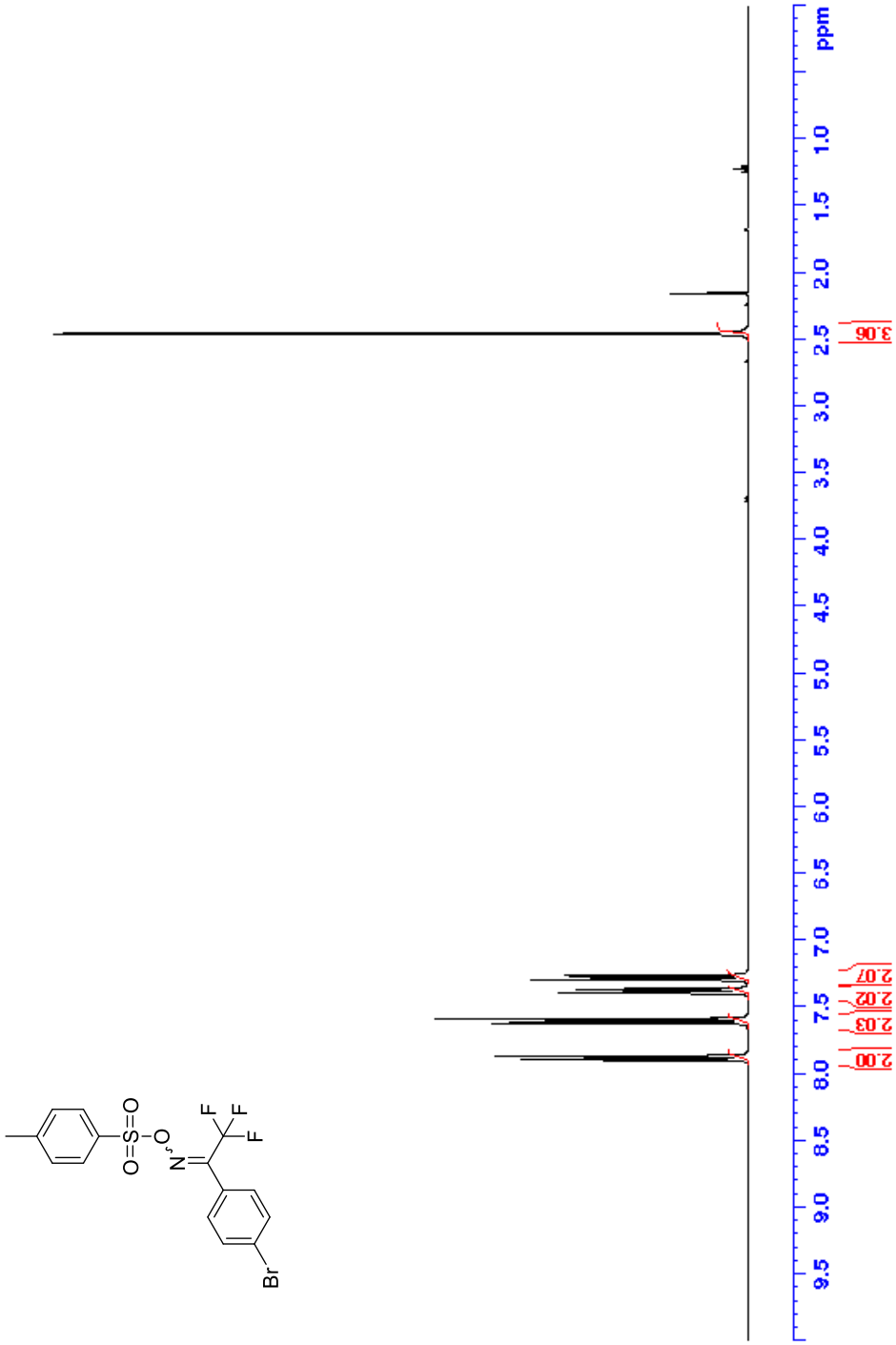
4'-bromo-2,2,2-trifluoroacetophenone (32) - ^{1}H NMR



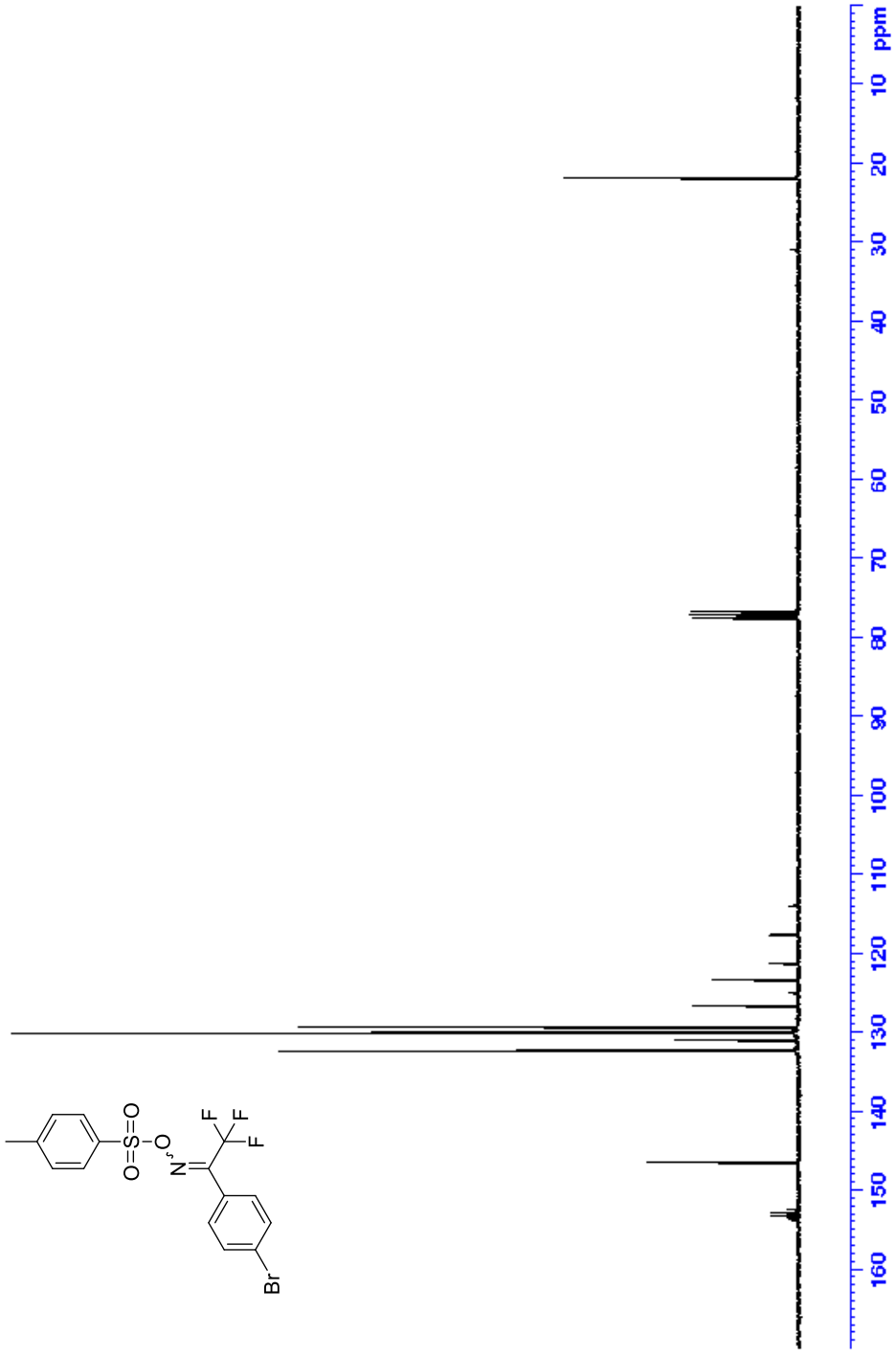
4'bromo-2,2,2-trifluoroacetophenone (32) - ^{13}C NMR



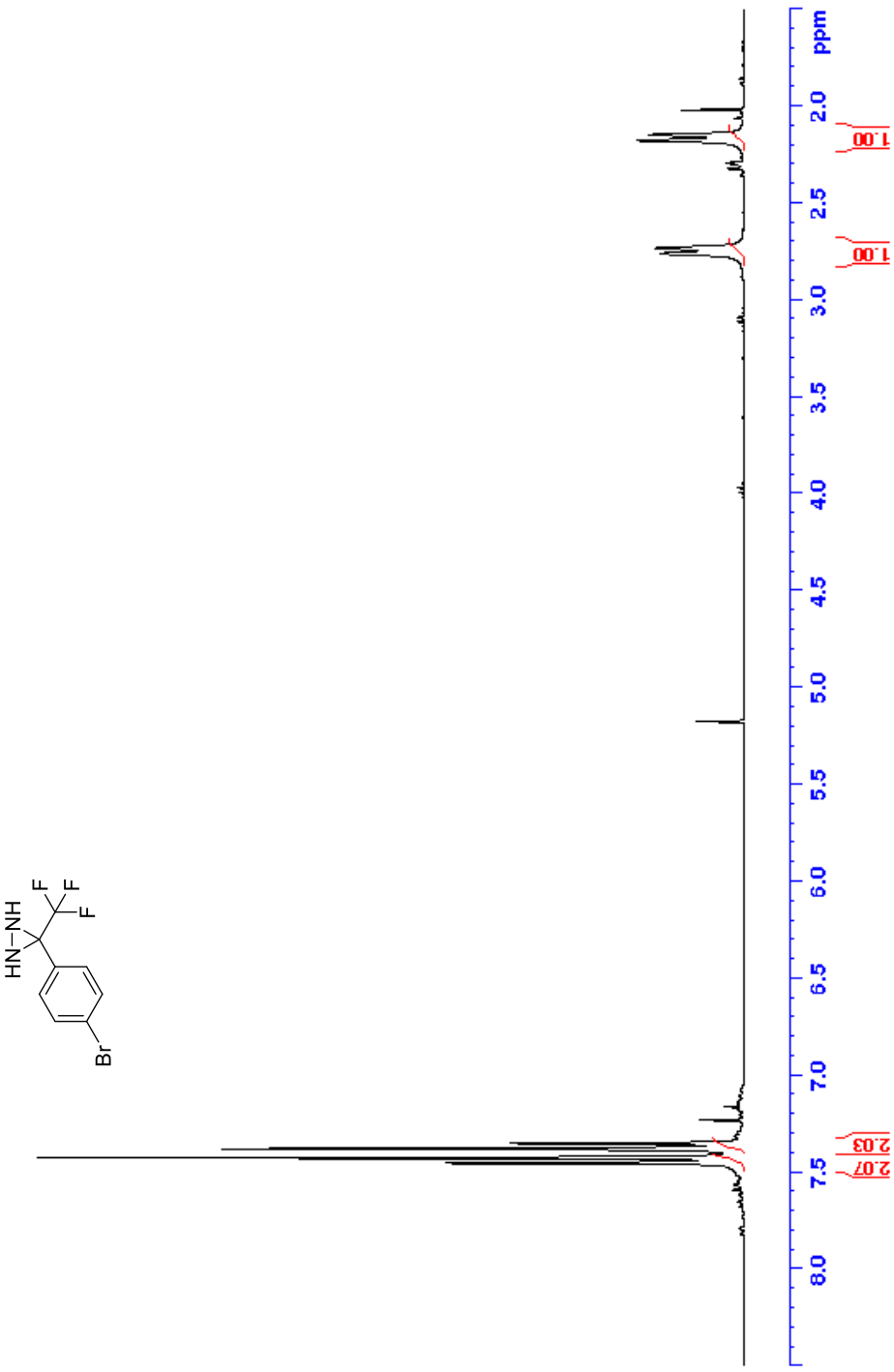
1-(4-bromophenyl)-2,2,2-trifluoroethanone *O*-tosyl oxime (34) - ^{1}H NMR



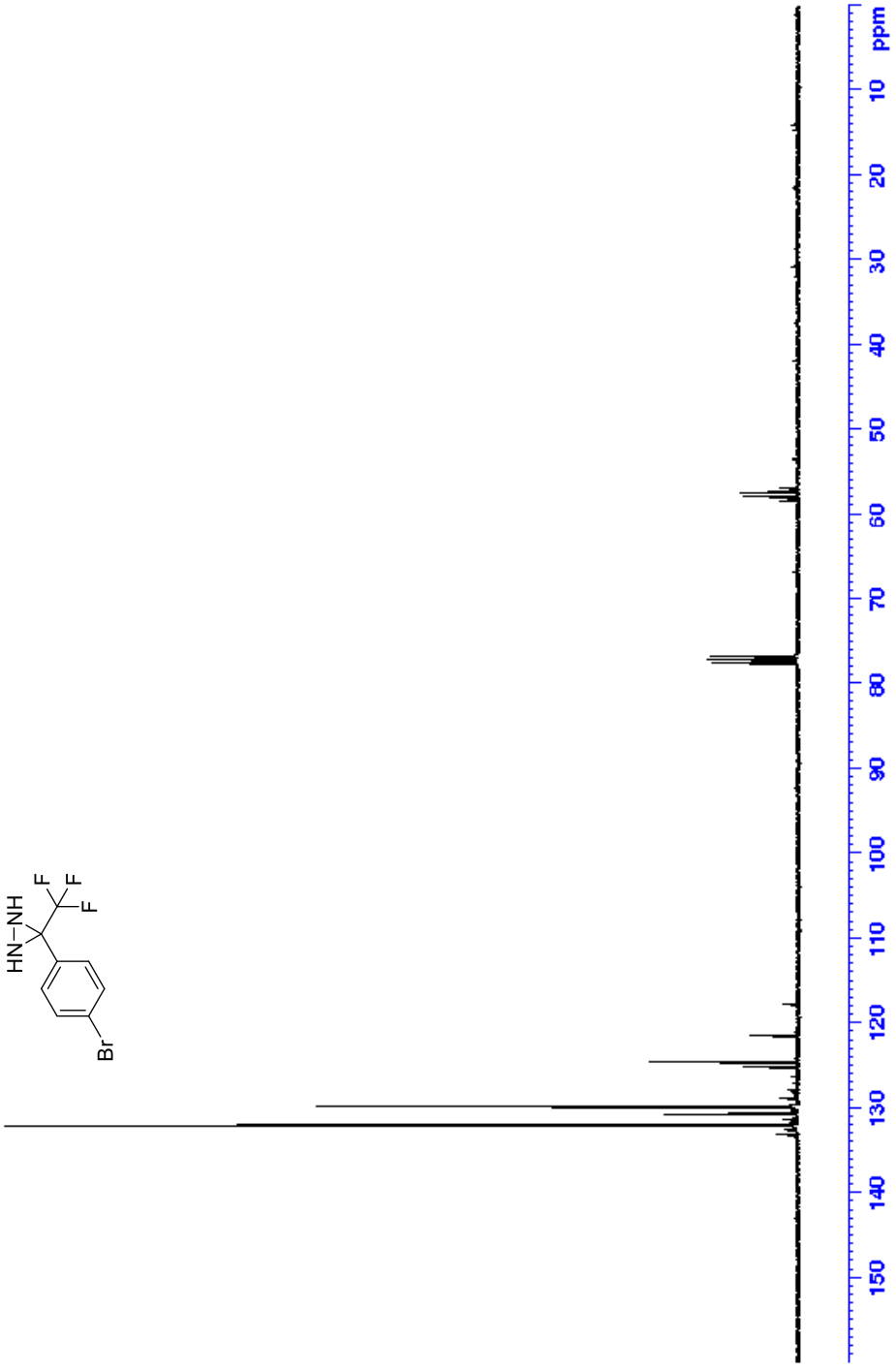
1-(4-bromophenyl)-2,2,2-trifluoroethanone *O*-tosyl oxime (34) - ^{13}C NMR



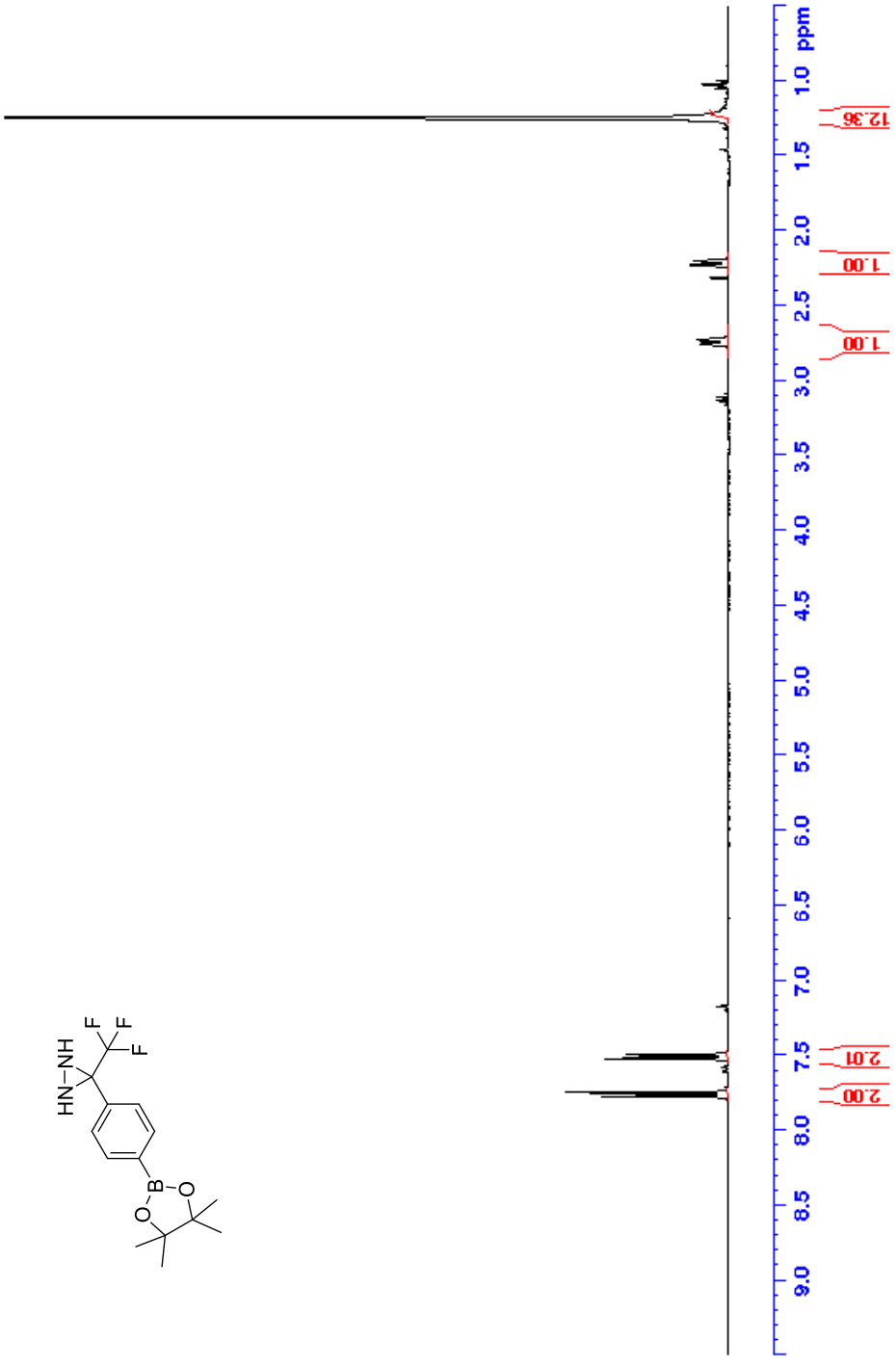
3-(4-bromophenyl)-3-(trifluoromethyl)diaziridine (35) - ^{1}H NMR



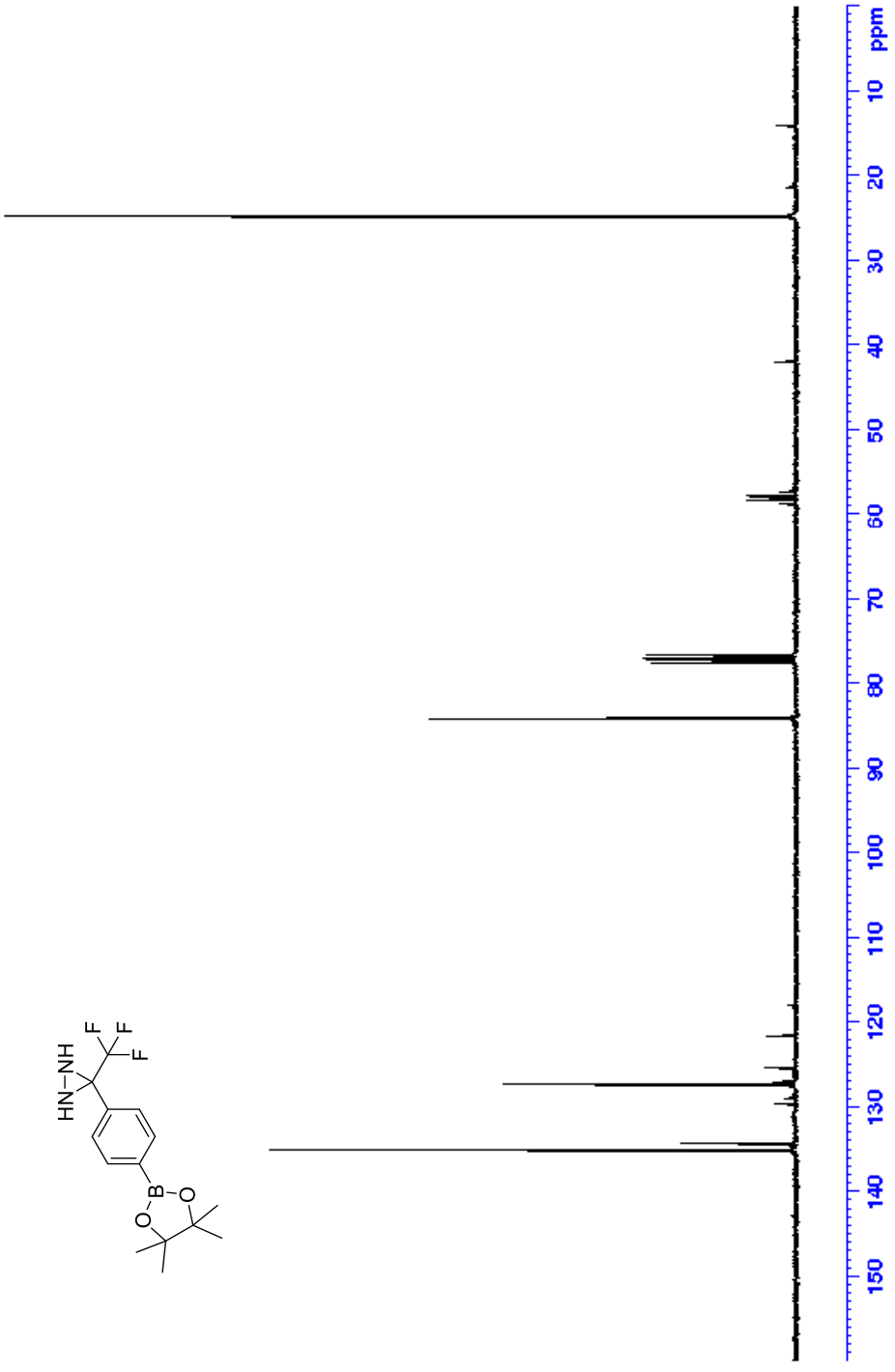
3-(4-bromophenyl)-3-(trifluoromethyl)diaziridine (35) - ^{13}C NMR



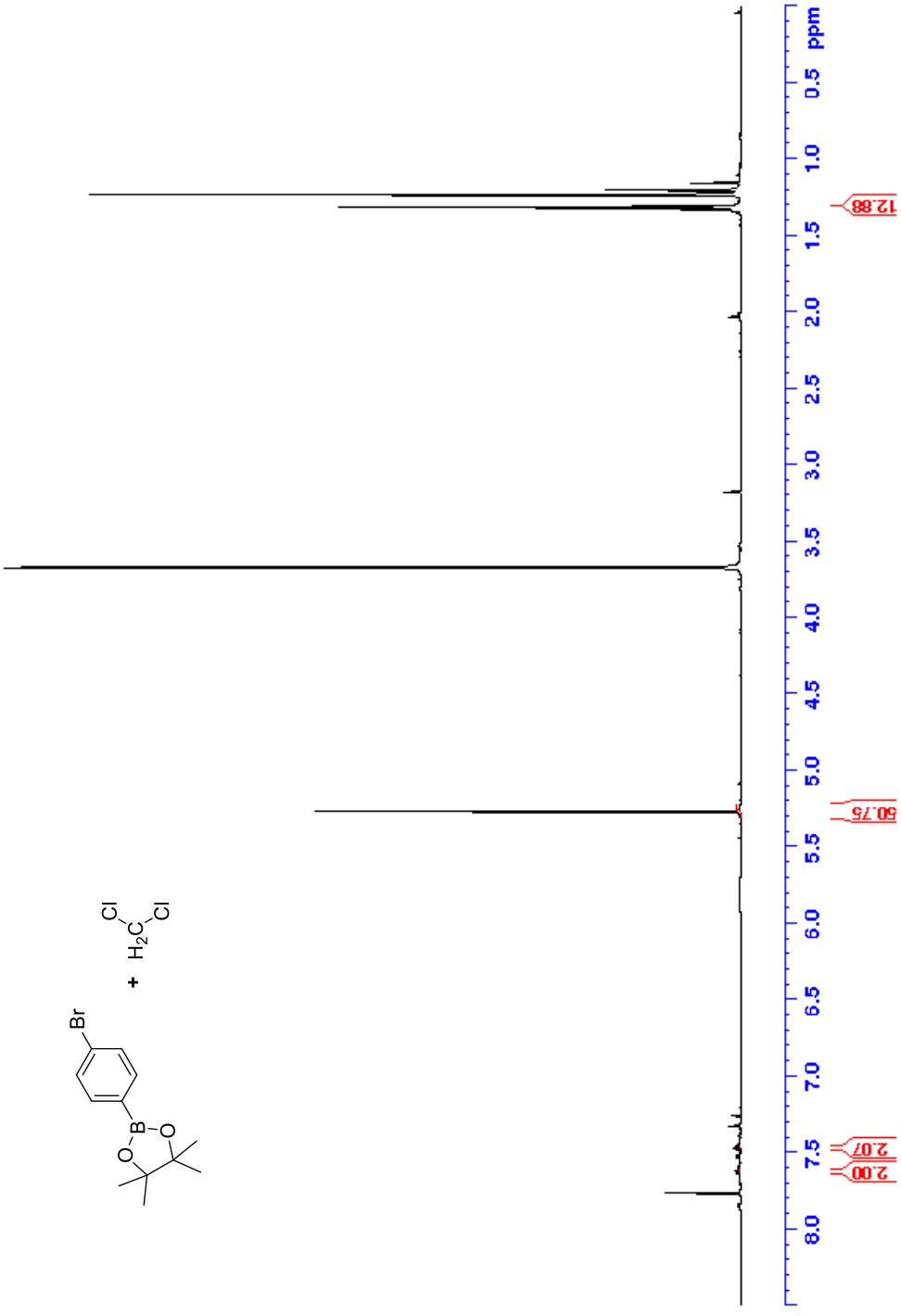
3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl) diaziridine (39) - ^{1}H NMR



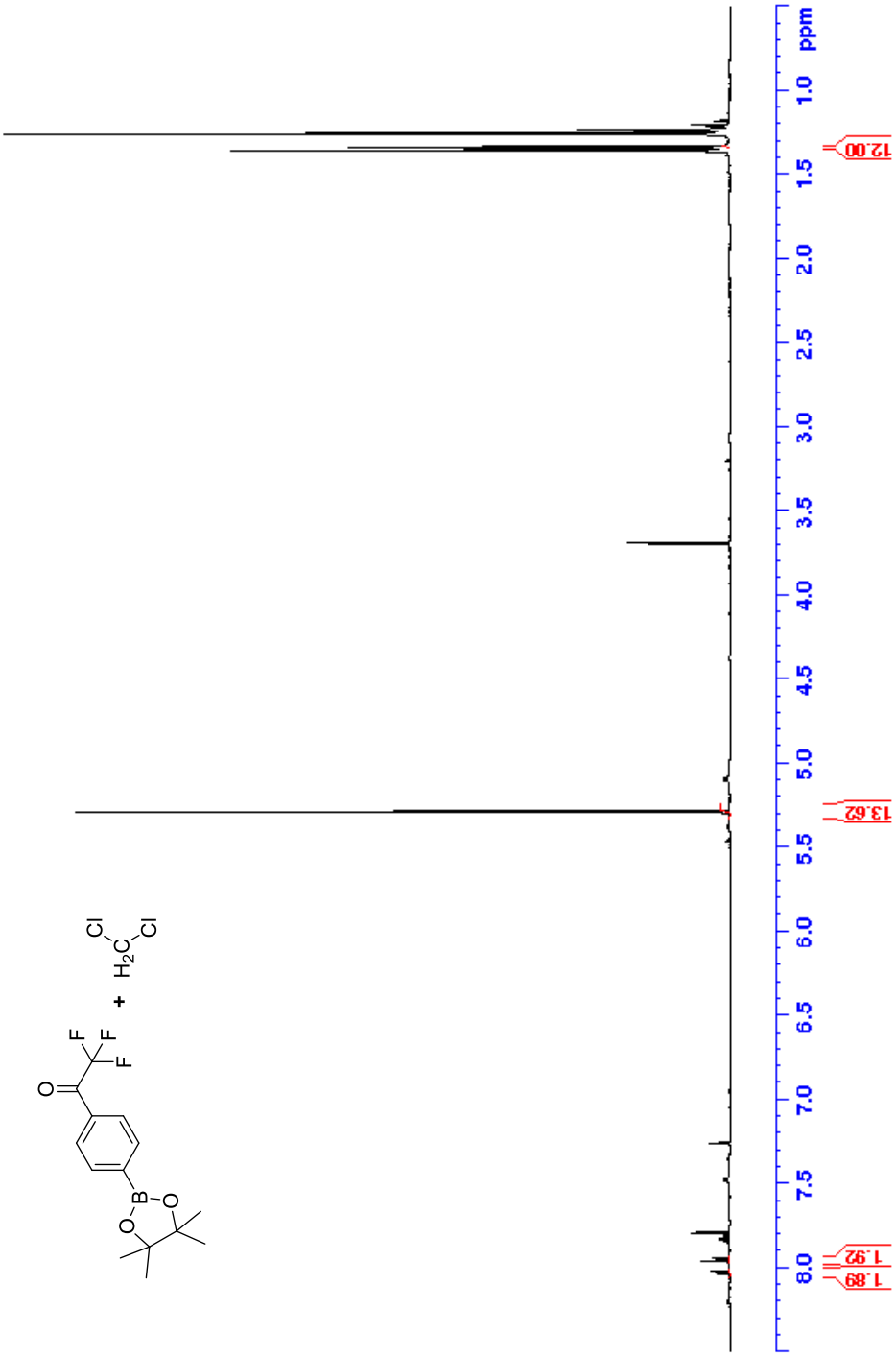
3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)-3-(trifluoromethyl) diaziridine (39) - ^{13}C NMR



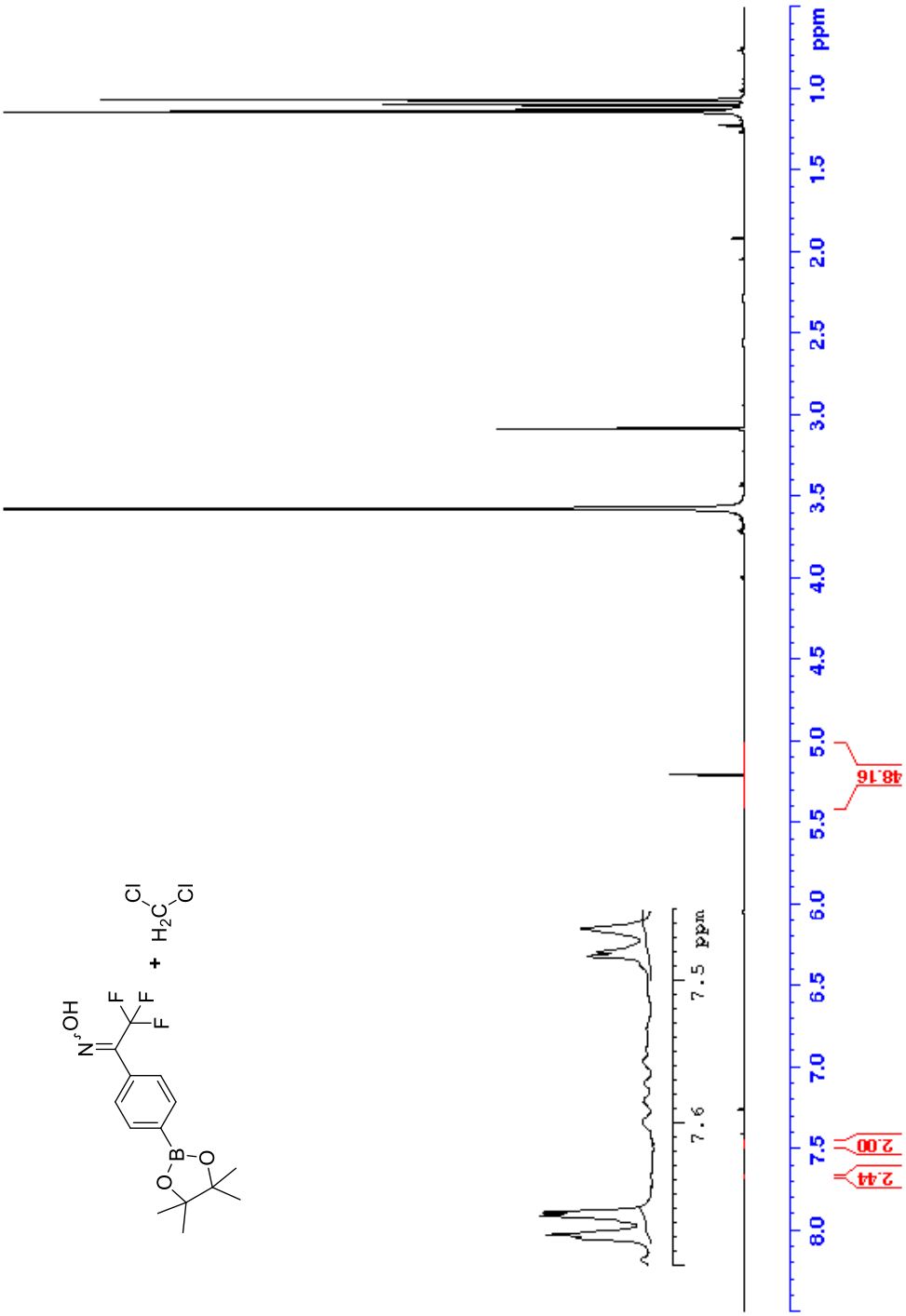
2-(4-bromophenyl)-4,4,5,5-tetramethyl-1,3,2-dioxaborolane (40) reaction mixture with DCM (6:1, DCM:theoretical yield) for conversion calculation - ^{1}H NMR



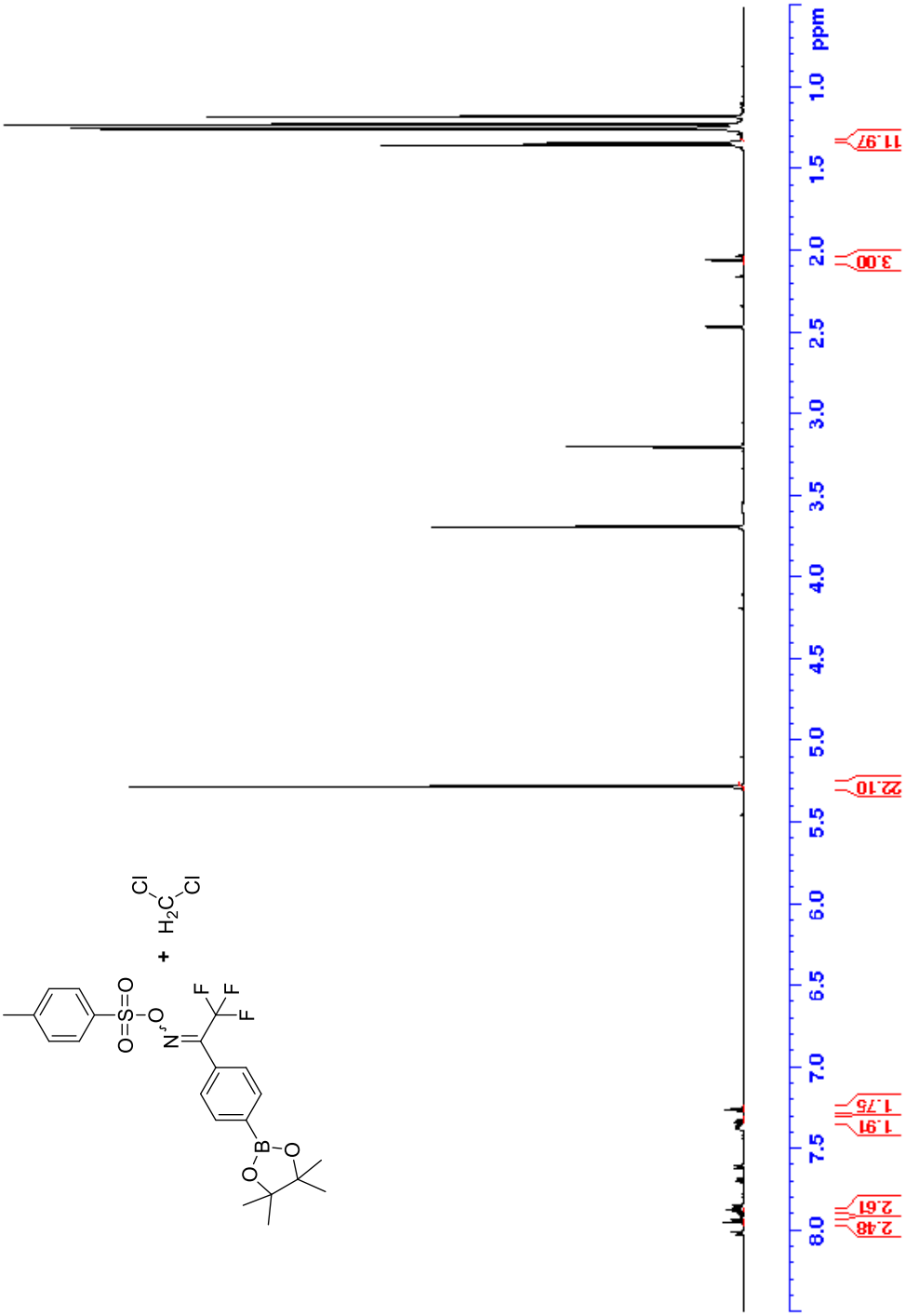
2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one (41) reaction mixture with DCM (6:1, DCM:theoretical yield) for conversion calculation - ^{1}H NMR



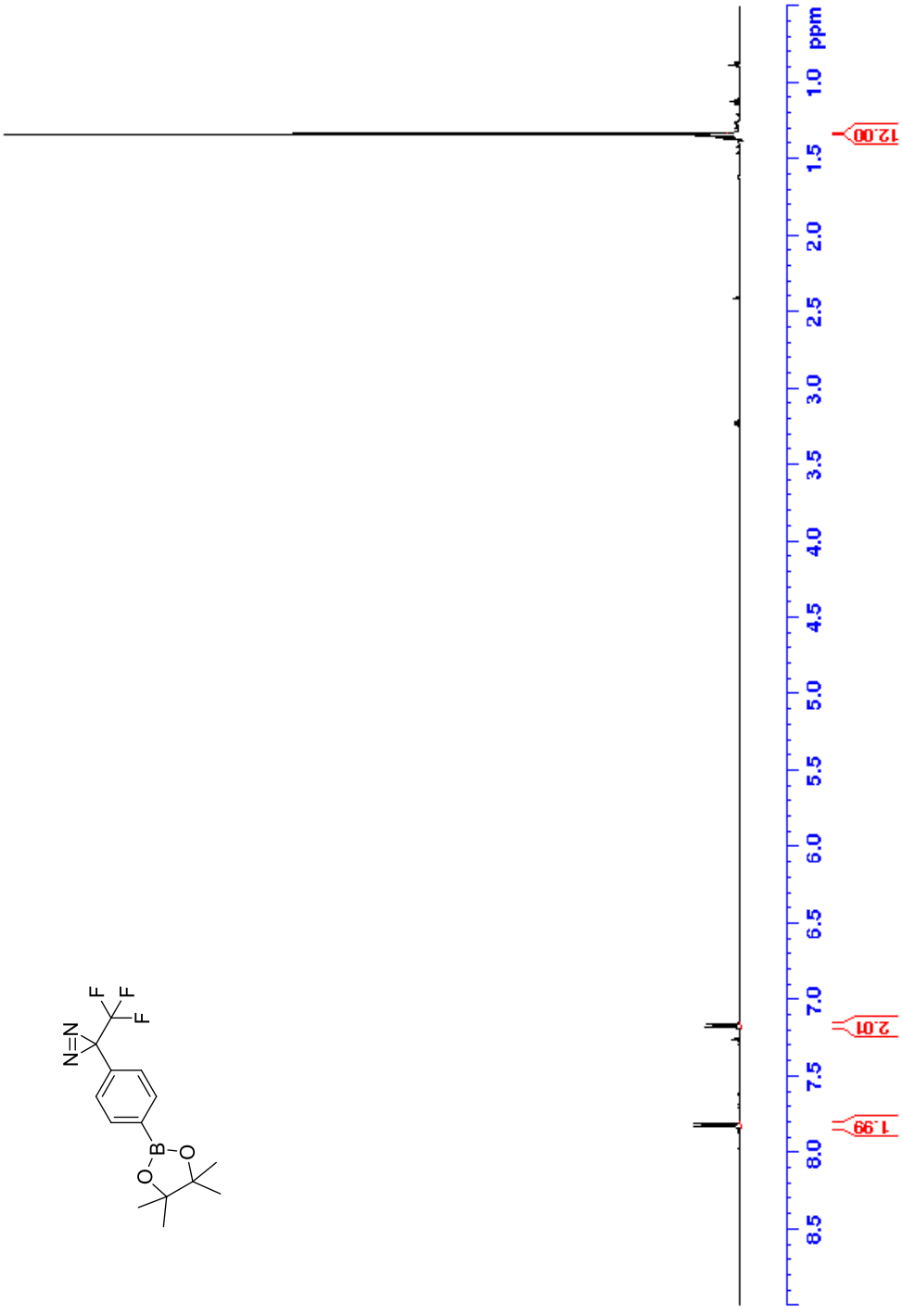
2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one oxime (42) reaction mixture with DCM (6:1, DCM:theoretical yield) for conversion calculation - ^{1}H NMR



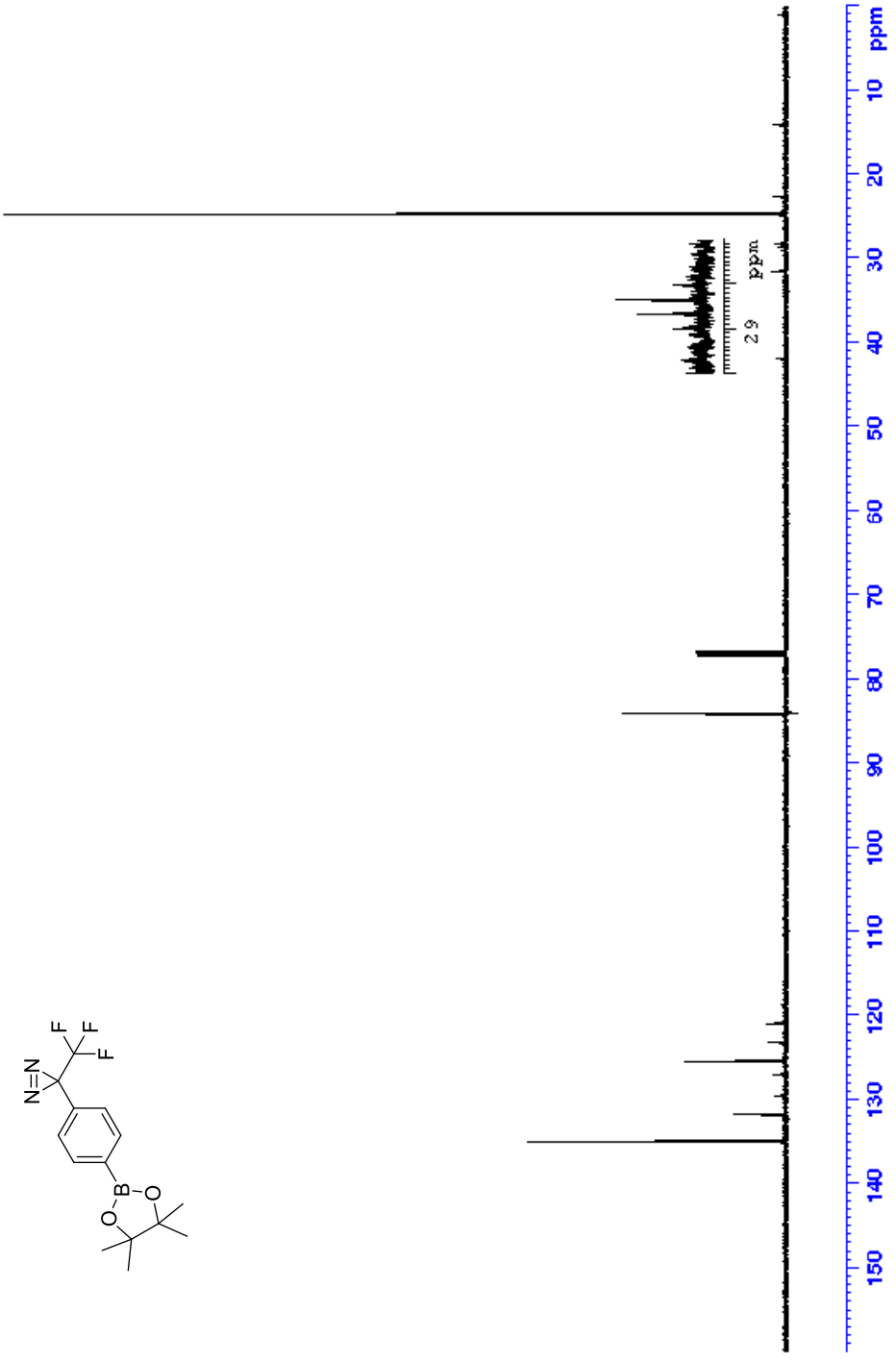
2,2,2-trifluoro-1-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl)ethan-1-one *O*-tosyl oxime (43) reaction mixture with DCM (6:1, DCM:theoretical yield) for conversion calculation - ^{1}H NMR



3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine (46) - ^{1}H NMR

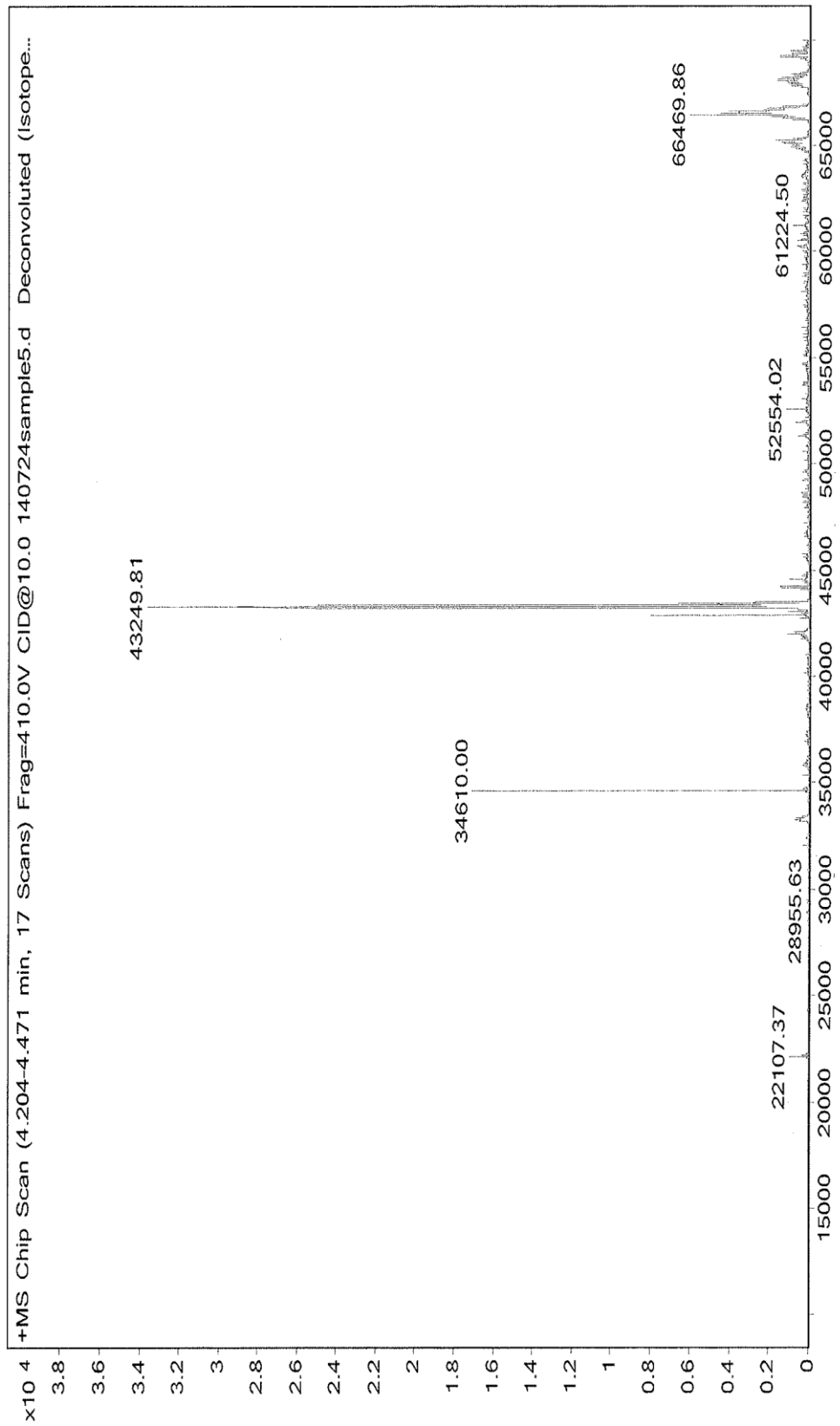


3-(4-(4,4,5,5-tetramethyl-1,3,2-dioxaborolan-2-yl)phenyl-3-(trifluoromethyl)-3*H*-diazirine (46) - ^{13}C NMR

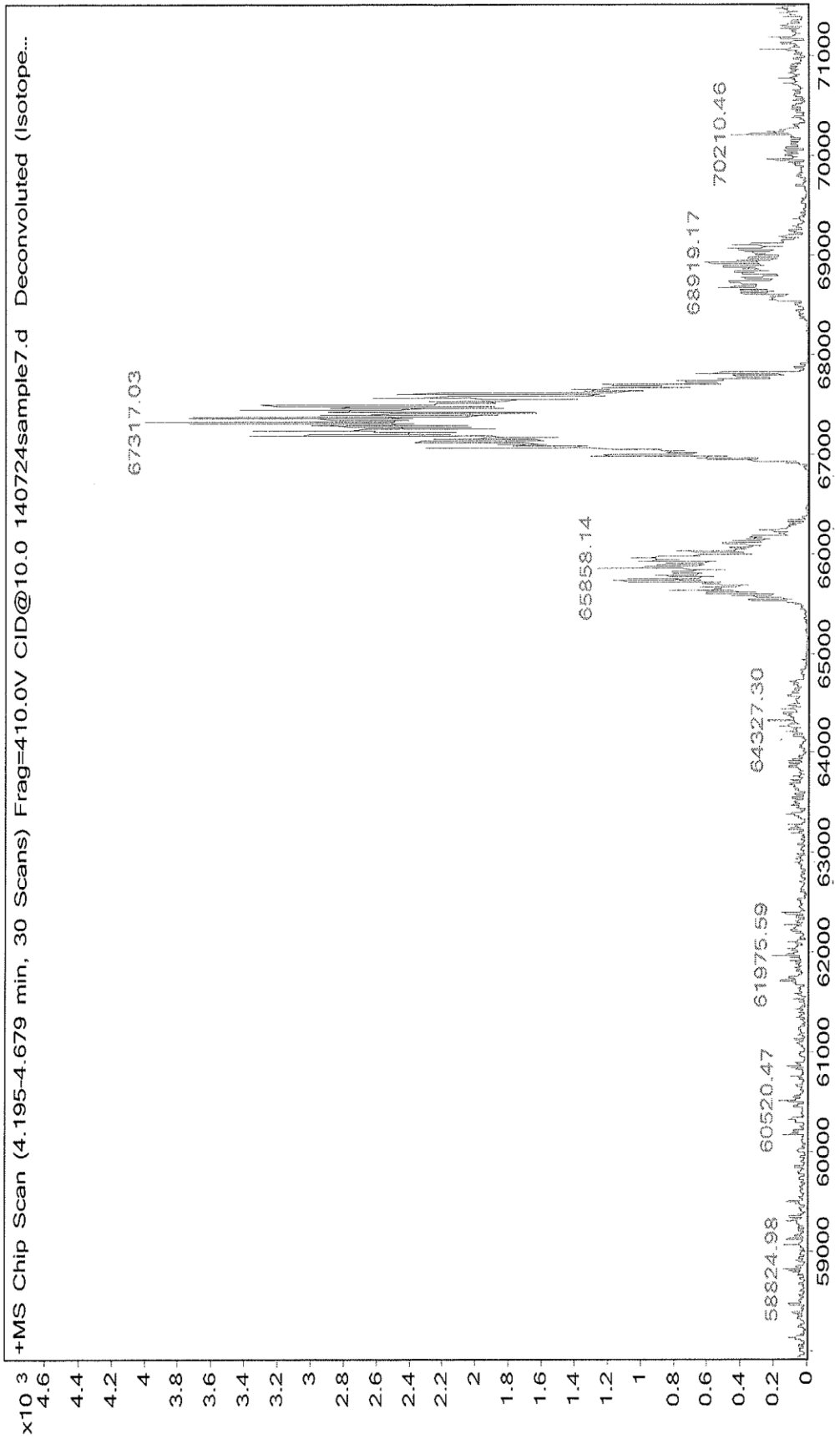


Appendix II: Mass Spectrometry Data

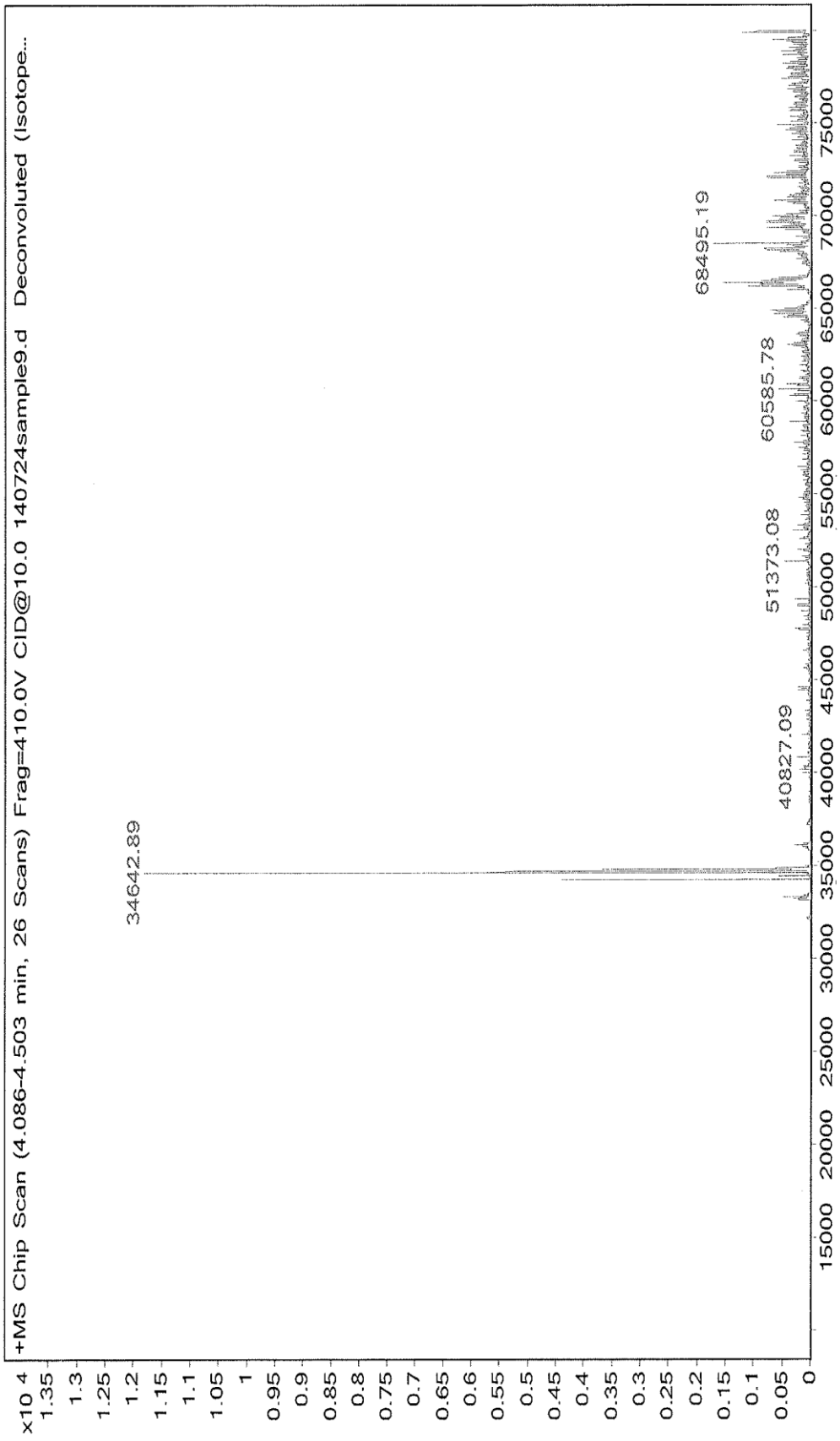
Human Serum Albumin



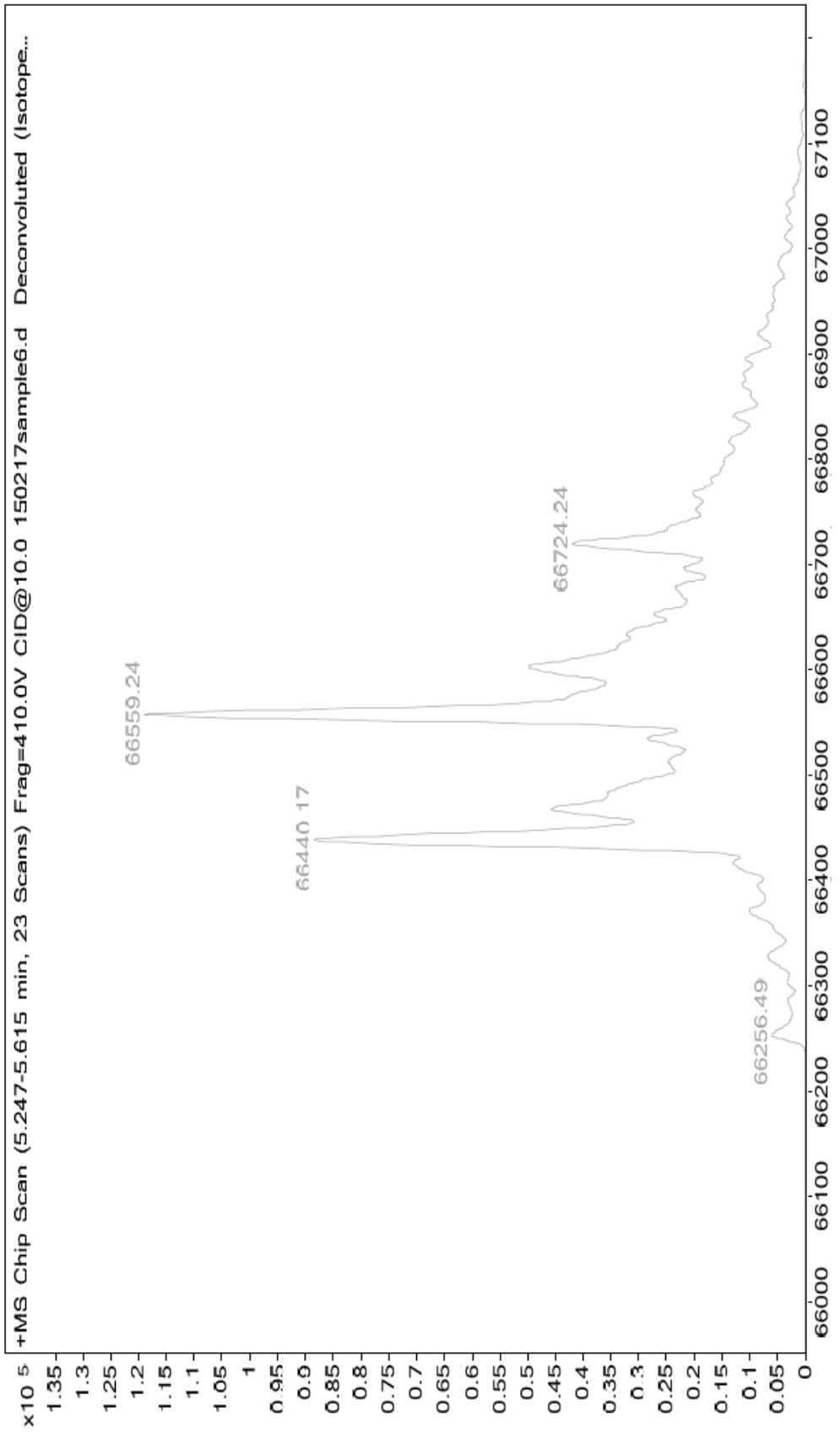
Human Serum Albumin labelled with 39 **I (5:1 probe:protein)**



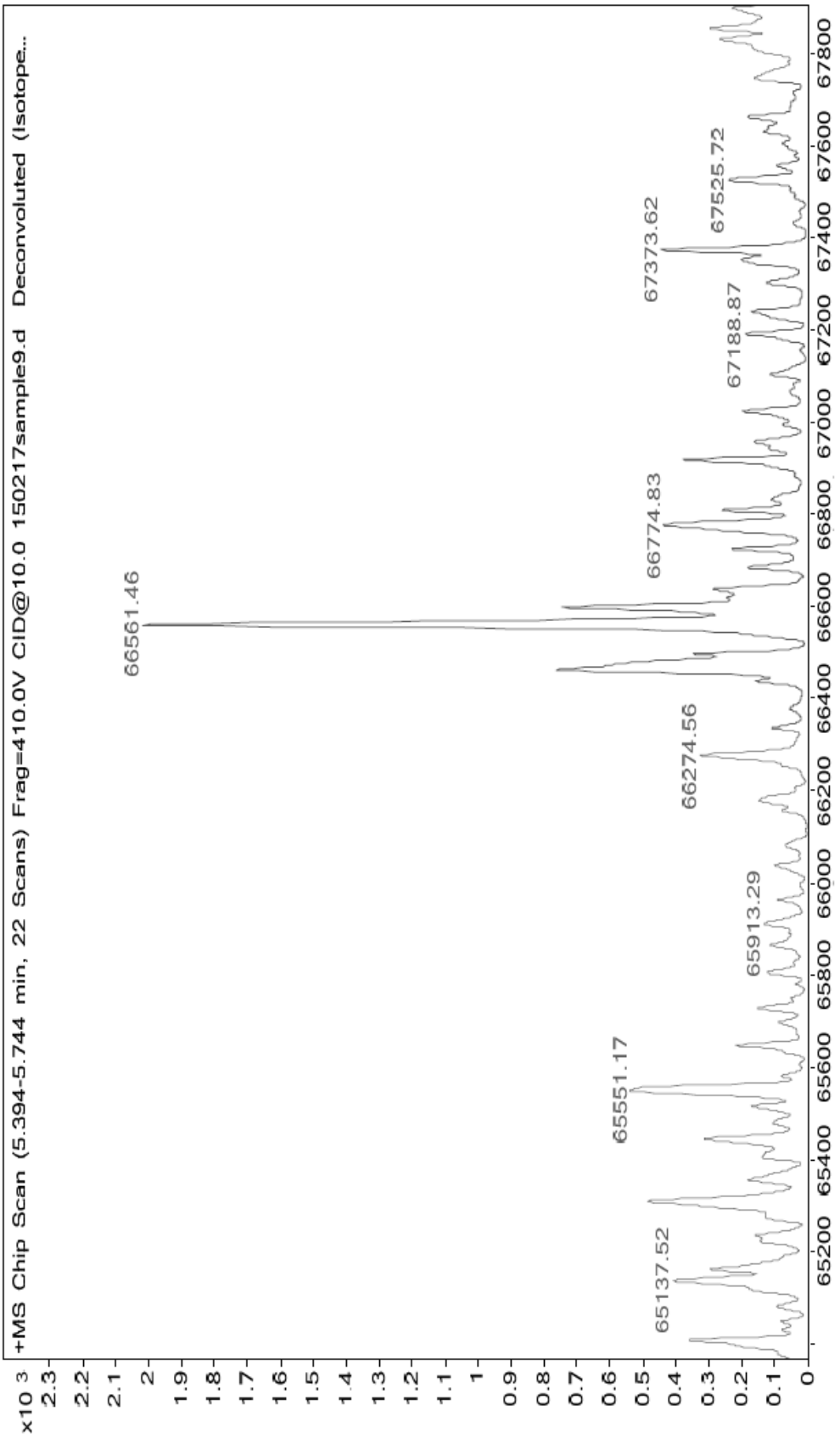
Human Serum Albumin labelled with 39 II (10:1 probe:protein)



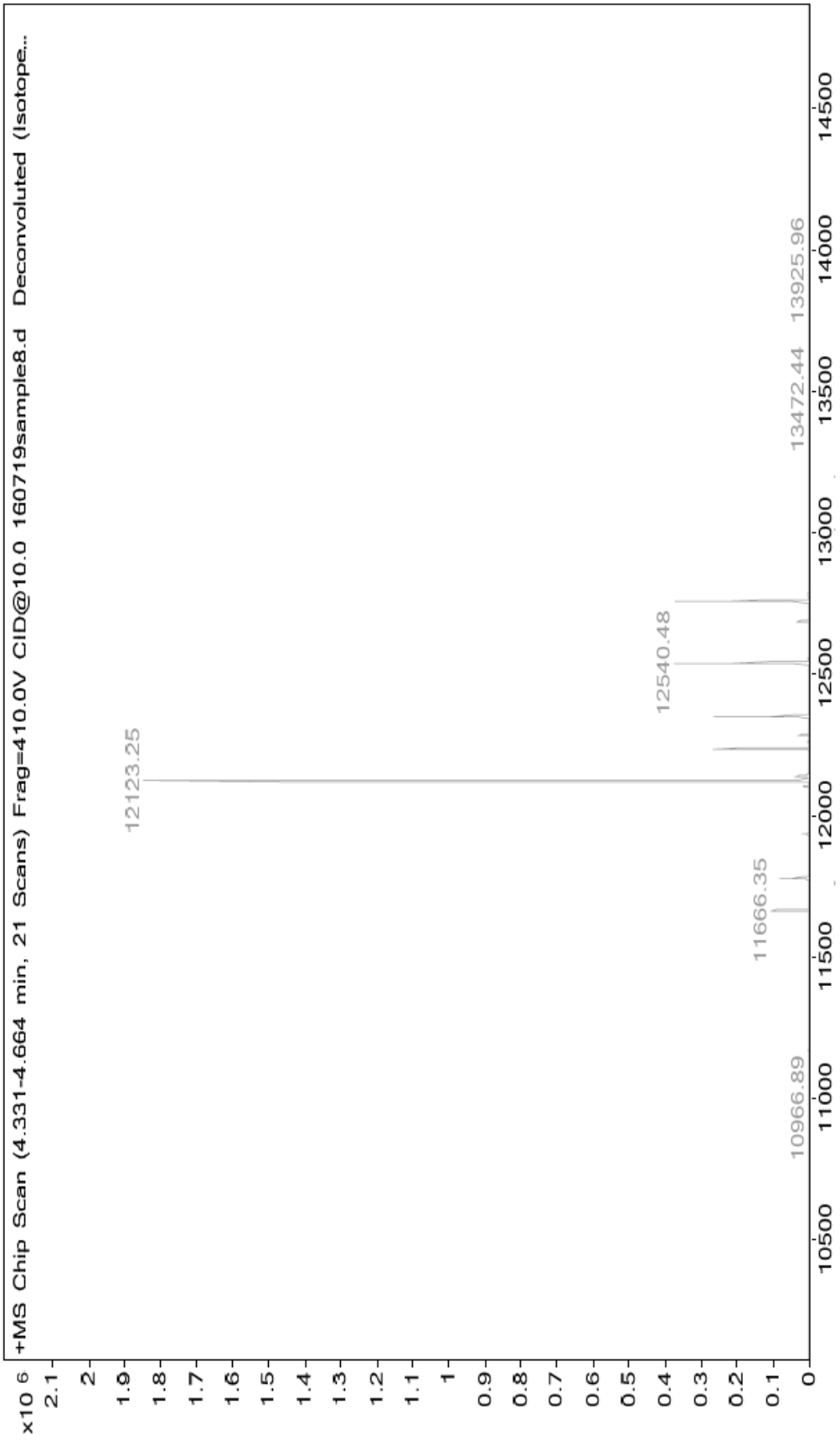
Human Serum Albumin Non-Targeting Control - 35 **used in place of probe (5:1 compound:protein)**



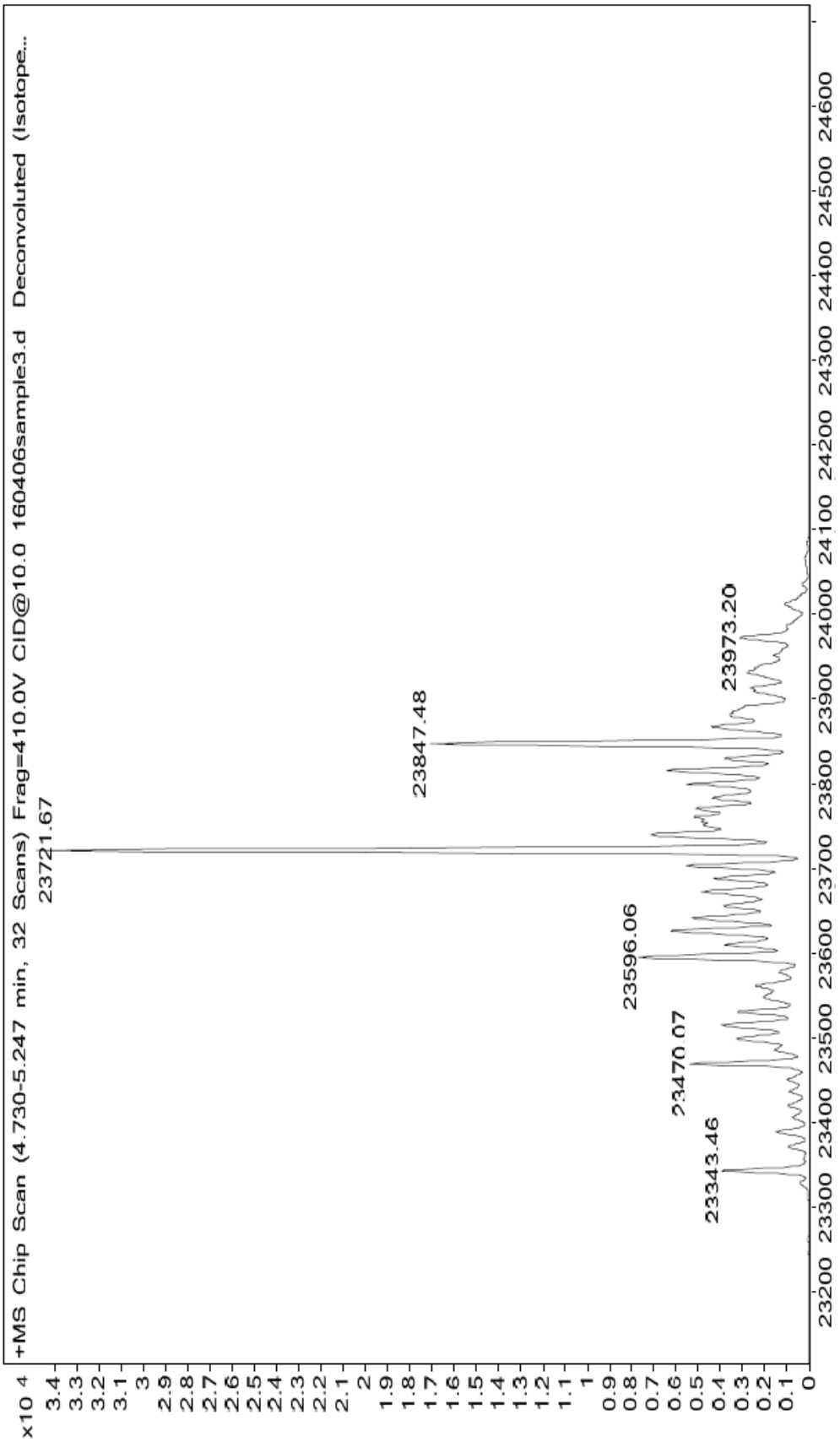
Human Serum Albumin Non-Coupling Control - 39 **as probe without oxidation step (5:1 compound:protein)**



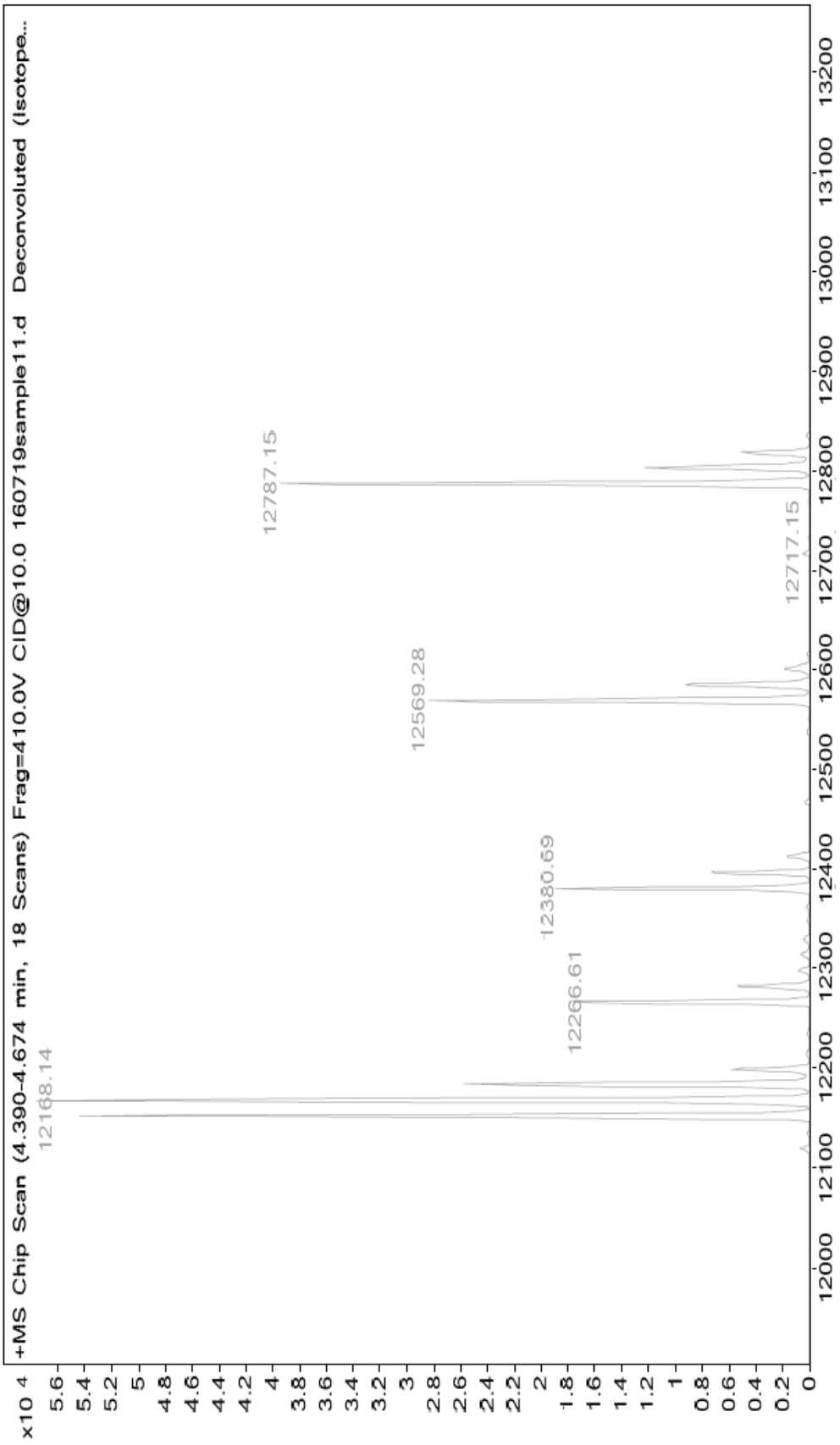
Glycated Macrophage Migration Inhibitory Factor



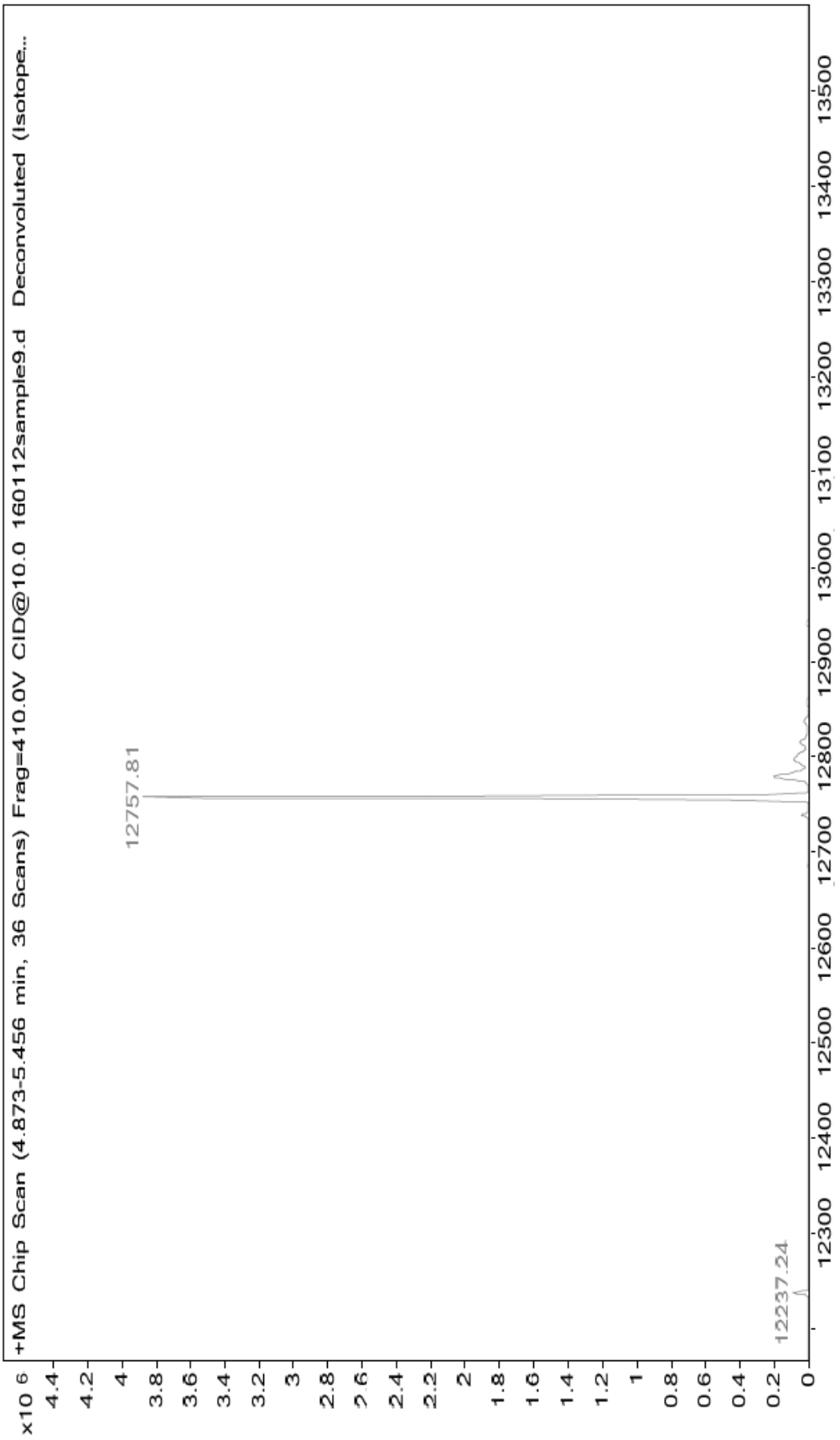
Glycated Macrophage Migration Inhibitory Factor labelled with 39 I (5:1 probe:protein)



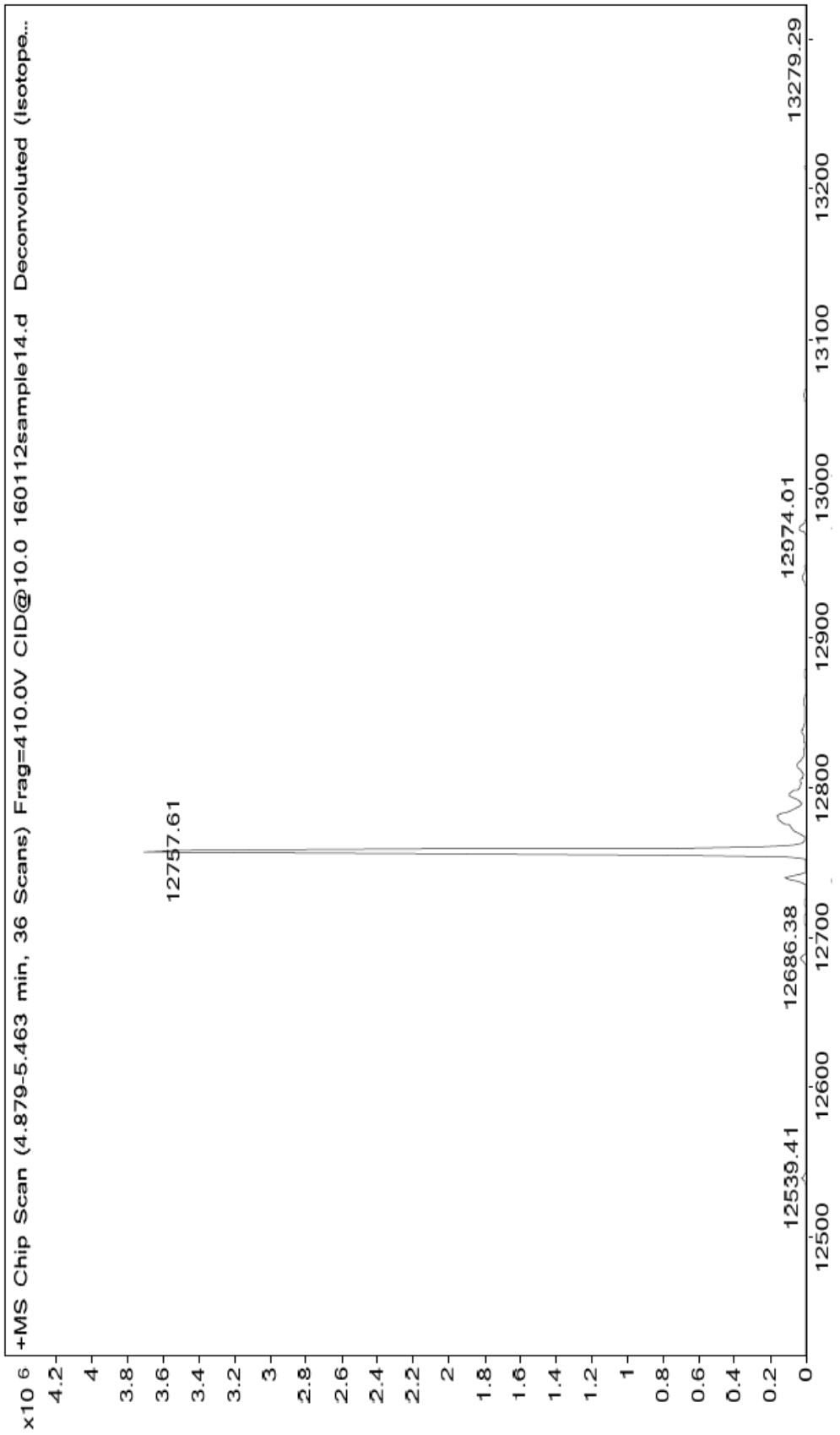
Glycated Macrophage Migration Inhibitory Factor labelled with 39 I I (5:1 probe:protein)



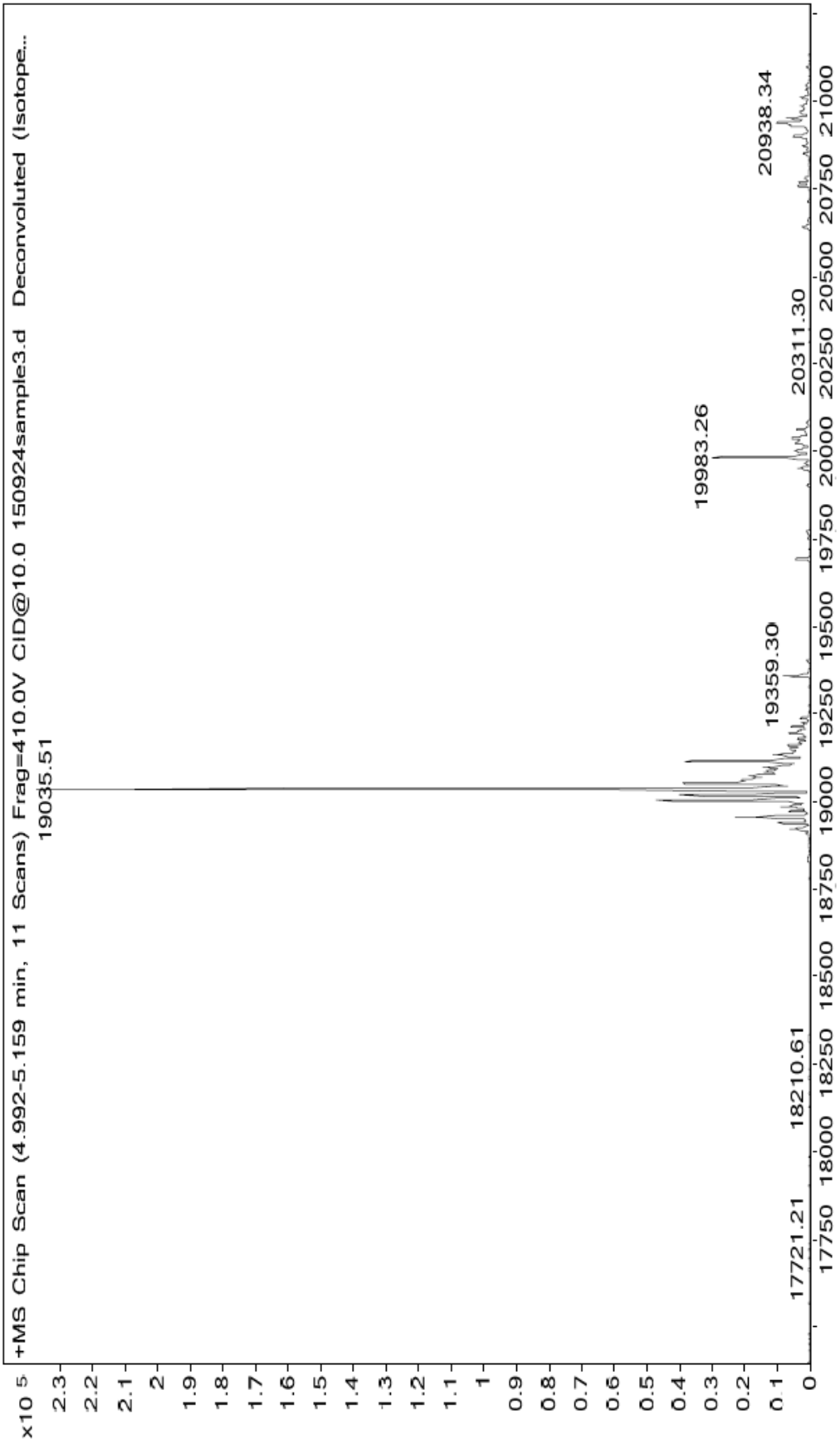
Non-Glycated Macrophage Migration Inhibitory Factor



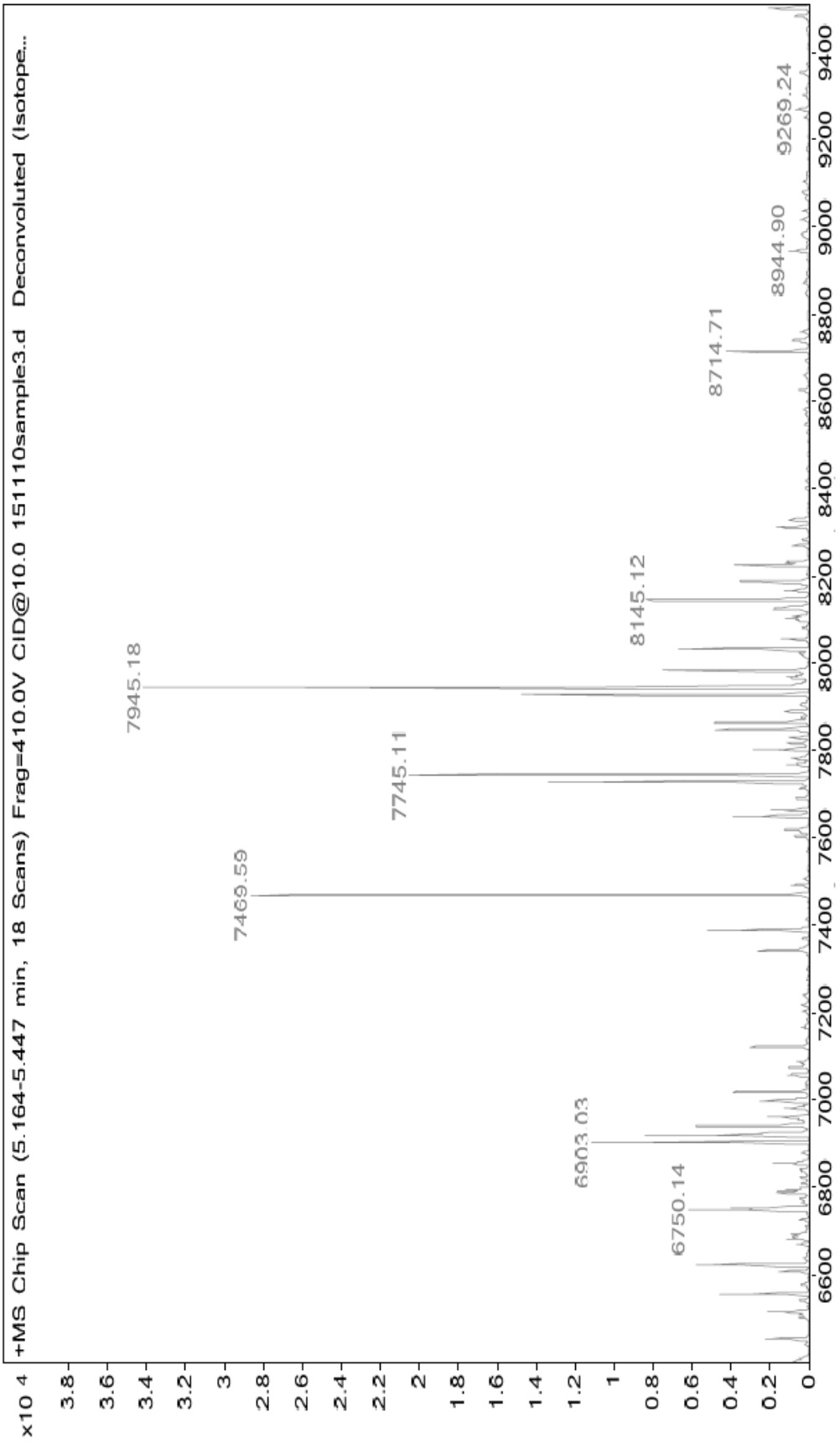
Non-Glycated Macrophage Migration Inhibitory Factor following labelling procedure - Non-Glycated Control



Casein


+MS Chip Scan (4.992-5.159 min, 11 Scans) Frag=410.0V CID@10.0 150924sample3.d Deconvoluted (Isotope...
19035.51
19983.26
20311.30
20938.34
19359.30
18210.61
17721.21
x10 5
2.3
2.2
2.1
2
1.9
1.8
1.7
1.6
1.5
1.4
1.3
1.2
1.1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
17750 18000 18250 18500 18750 19000 19250 19500 19750 20000 20250 20500 20750 21000

Casein labelled with 39 I (5:1 probe:protein)



Casein Labelled with 39 II (5:1 probe:protein)